This is a digital copy of a book that was preserved for generations on library shelves before it was carefully scanned by Google as part of a project to make the world's books discoverable online.

It has survived long enough for the copyright to expire and the book to enter the public domain. A public domain book is one that was never subject to copyright or whose legal copyright term has expired. Whether a book is in the public domain may vary country to country. Public domain books are our gateways to the past, representing a wealth of history, culture and knowledge that's often difficult to discover.

Marks, notations and other marginalia present in the original volume will appear in this file - a reminder of this book's long journey from the publisher to a library and finally to you.

Usage guidelines

Google is proud to partner with libraries to digitize public domain materials and make them widely accessible. Public domain books belong to the public and we are merely their custodians. Nevertheless, this work is expensive, so in order to keep providing this resource, we have taken steps to prevent abuse by commercial parties, including placing technical restrictions on automated querying.

We also ask that you:

+ *Make non-commercial use of the files* We designed Google Book Search for use by individuals, and we request that you use these files for personal, non-commercial purposes.

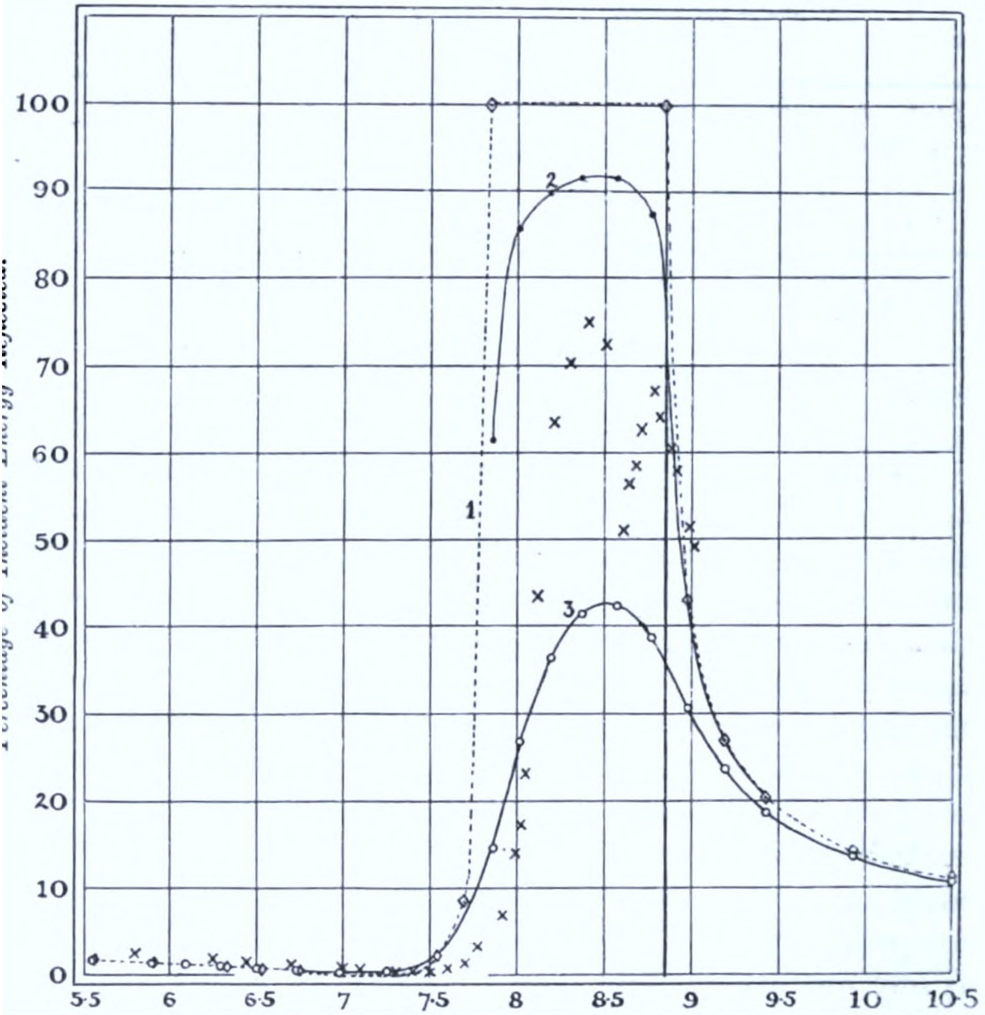
+ *Refrain from automated querying* Do not send automated queries of any sort to Google's system: If you are conducting research on machine translation, optical character recognition or other areas where access to a large amount of text is helpful, please contact us. We encourage the use of public domain materials for these purposes and may be able to help.

+ *Maintain attribution* The Google "watermark" you see on each file is essential for informing people about this project and helping them find additional materials through Google Book Search. Please do not remove it.

+ *Keep it legal* Whatever your use, remember that you are responsible for ensuring that what you are doing is legal. Do not assume that just because we believe a book is in the public domain for users in the United States, that the work is also in the public domain for users in other countries. Whether a book is still in copyright varies from country to country, and we can't offer guidance on whether any specific use of any specific book is allowed. Please do not assume that a book's appearance in Google Book Search means it can be used in any manner anywhere in the world. Copyright infringement liability can be quite severe.

About Google Book Search

Google's mission is to organize the world's information and to make it universally accessible and useful. Google Book Search helps readers discover the world's books while helping authors and publishers reach new audiences. You can search through the full text of this book on the web at `http://books.google.com/`



The electron theory of matter

Owen Willans Richardson

THE
ELECTRON THEORY
OF
MATTER

BY

O. W. RICHARDSON

M.A. (CAMB.), D.Sc. (LOND.), F.R.S.

WHEATSTONE PROFESSOR OF PHYSICS AT KING'S COLLEGE, LONDON

Cambridge:
at the University Press
1914

Cambridge:
PRINTED BY JOHN CLAY, M.A.
AT THE UNIVERSITY PRESS

UNIV. OF
CALIFORNIA

CAMBRIDGE UNIVERSITY PRESS
C. F. CLAY, MANAGER
London: FETTER LANE, E.C.
Edinburgh: 100 PRINCES STREET

London: H. K. LEWIS, 136 GOWER STREET, W.C.
London: WILLIAM WESLEY AND SON, 28 ESSEX STREET, STRAND
Berlin: A. ASHER AND CO.
Leipzig: F. A. BROCKHAUS
New York: G. P. PUTNAM'S SONS
Bombay and Calcutta: MACMILLAN AND CO., LTD.
Toronto: J. M. DENT AND SONS, LTD.
Tokyo: THE MARUZEN-KABUSHIKI-KAISHA

All rights reserved

PREFACE

THIS book is based on a course of lectures which I have delivered to graduate students at Princeton University during the last few years. My aim has been to exhibit the extent to which the fundamental facts of physical science may be coordinated by means of the conception of the electron and the laws of electrodynamics. In developing the subject I have started from the most elementary beginnings, and I have therefore found it necessary to include much matter which is to be found in any ordinary text-book of the theory of electricity and magnetism. It is hoped that the lack of conciseness thereby involved may be more than atoned for by the wider circle to which the book may appeal. The course of lectures at Princeton on which the book is founded proved useful as an introduction to the methods of modern mathematical physics in addition to forming a presentation of the results of recent physical discovery.

The broad scope of the subject makes it imperative that a good deal of selection should be exercised as to the nature and treatment of the topics considered. In determining these, consideration has been given to importance, interest, and instructiveness, roughly in the order named. The necessary incompleteness is remedied to some extent by references to scientific papers and to other works. These references are intended to supplement the discussion in the text rather than to exhibit the historical development of the subject. Thus many important papers are not referred to. I have tried, however, to be as accurate as possible in any statements which deal specifically with historical matters.

For a variety of reasons the book has, unfortunately, suffered considerable delay in passing through the press. I have, however, found it possible to incorporate some account of the important recent results while correcting the proofs; so that, with some reservations, the book may be regarded as fairly representing the

CHAPTER I

THE ORIGIN OF THE ELECTRON THEORY

The Electron Theory of Matter may be looked upon as a form of atomic theory. It differs from the form of the atomic theory with which chemistry is familiar, especially in that it makes the ultimate atoms minute geometrical configurations of electric charge, instead of particles of uncharged matter. A large number of different lines of inquiry, often closely interwoven, have led to the adoption of such a view of the structure of matter. Of these different lines of inquiry, however, three may be considered pre-eminently conspicuous.

In the first place, although the electron theory has made most rapid progress in the last two decades, it is a logical development of the views held a century ago by Davy* and Berzelius† and especially of the views to which Faraday‡ was led by his electrochemical discoveries made somewhat later. Davy concluded, from a general review of the electrochemical phenomena known in his day, that the forces between the chemical atoms were of electrical origin. Shortly afterwards a complete system of chemical structure depending on the same idea was developed by Berzelius and, although in its original form Berzelius's electrochemical theory was insufficiently elastic, its main features have much in common with the most modern views on the subject. The laws of electrolysis discovered by Faraday led to an important advance by pointing distinctly to an atomic constitution for electricity; for they showed that each chemical atom invariably transported either a definite quantity of electricity or an integral multiple of that quantity, a multiple which was determined by the chemical valency of the atom. This inference from Faraday's electrolytic researches was strongly advocated much later by von Helmholtz*.

---

* *Phil. Trans.* p. 1 (1807).

† *Mem. Acad. Stockholm* (1812); *Nicholson's Journal*, vols. xxxiv. and xxxv. (1813).

‡ *Exp. Res.* §§ 377, 523, 661, 713, 821 and especially 852, 869.

The second line of inquiry referred to dealt mainly with optical phenomena. It would be impossible adequately to discuss at this stage the complex questions which present themselves in this connection. It may be permissible to recall that Maxwell's electromagnetic theory had been found to account satisfactorily for the behaviour of light, and electromagnetic waves in general, in free space; but that difficulties presented themselves when phenomena like refraction and dispersion, which depend upon transmission through material media, were considered. These were found capable of removal by introducing the simple hypothesis that the material media contained particles having appropriate natural frequencies of vibration. Maxwell's electromagnetic theory of light naturally suggested that these particles were electrically charged, and the facts of dielectric polarisation were then found to fall into line, approximately at any rate, with the optical phenomena. The theory of the propagation of light in moving media also made important advances under the influence of the new views. A striking confirmation of the correctness of the general position was furnished by Zeeman's discovery of the change of the frequency of spectral lines when the emitting source was placed in a strong magnetic field. The magnitude of this change enabled an estimate of the ratio of the charge to the mass of the particles to be made. The resulting values were in substantial agreement with those which were obtained at about the same time by entirely different and more direct methods. In the development of the theory along the lines just indicated the ideas of H. A. Lorentz† and J. Larmor‡ have had a preponderating influence.

The third line of attack was furnished by the experimental study of the phenomena accompanying the discharge of electricity through gases and especially of the properties of the cathode rays, the Roentgen rays and the rays emitted by radioactive substances. The matter described in the ensuing paragraph will serve to give total charge. Stated briefly the method is as follows: The amount of water condensed on the whole of the drops in the cloud may be calculated from the degree of supersaturation produced by the known expansion. The application of Stokes's formula for the rate of fall of a sphere in a viscous fluid gives the average size of each drop. Thus these data determine the total number n of the drops. It is assumed, for sufficient reasons, that practically all the drops contain one and only one ion. The total charge ne on all the ions could be determined by sweeping them out of the chamber into an electrometer before the expansion took place, the strength of the source of ionisation being the same as in the condensation experiments. Thus the charge on a single ion was obtained by division. Proceeding in this way Thomson showed that the negative ions liberated in air by Roentgen rays and by the β rays from radium each carried the same charge as the hydrogen ion in electrolysis.

---

* "Faraday Lecture" (1881).

† *Arch. Neerl.* vol. xxv. p. 363 (1892). *Theory of Electrons*, Leipzig, 1909.

‡ *Phil. Trans.* A. vol. CLXXXV. p. 821 (1894). *Aether and Matter*, Cambridge, 1900.

The ions investigated in these experiments are rather complicated structures and are not identical with the electron. In the case of the ionisation produced by ultra-violet light falling on a metal it was shown by Thomson that the particles when first emitted have the same value of e/m as the cathode rays. These would not be likely to aggregate together in the presence of gas molecules, and C. T. R. Wilson* showed that the negative ions from ultra-violet light behaved exactly like those from the other ionising agents in his condensation experiments. The inference from these experiments therefore is that the particles which form the cathode rays and which are emitted during photoelectric action carry a charge equal to that of the hydrogen atom in electrolysis. Experiments by Townsend†, on the rate of fall of the clouds produced when the gases evolved from chemical actions occurring in the wet way are allowed to bubble through water, had previously led him to conclude that the ions present in such gases carry the same charge as a hydrogen ion in electrolysis.

This conclusion has been strengthened by other methods of determining the charge on an electron. One of these depends on the theory of the radiation of electromagnetic energy from hot bodies. The theory of this method will be considered in the sequel*. One of the recent methods, which is due to Rutherford† and Geiger, depends upon the properties of radioactive substances. These are found to emit positively charged bodies, called α particles, which carry twice the charge e of an electron and are able to produce a large number of new ions when they pass through a gas. By magnifying this secondary ionisation by means of an auxiliary electric field and also using a very sensitive electrometer, Rutherford was able to detect the ionisation produced by a single α particle. When a very weak radioactive preparation was used the α particles were emitted at times separated by rather wide and irregular intervals, and as the effect produced by each one separately could be detected, the number emitted by a given amount of the radioactive substance in a given time could be measured. The only other datum which is required to measure e is the quantity of positive electricity which is carried away from the same quantity of the preparation by the α rays. This had previously been obtained by other experiments.

---

* *Phil. Trans.* A. vol. CXCII. p. 403 (1899).

† *Phil. Mag.* Feb. 1898.

The falling drop method has recently been improved by H. A. Wilson‡ and R. A. Millikan§. The former showed that the charge on the drops could be deduced from the rate of fall under gravitation combined with different electric fields, without making use of the degree of supersaturation; whilst the latter showed how the drops of water could be replaced by drops of a non-volatile oil. The drops of oil have the great advantage that they do not evaporate: and by allowing a sufficient number of electrons to combine with them and applying a supporting electric field which just balances the gravitational force, they can be kept under observation for an indefinite length of time. In this way Millikan has shown that the method is capable of yielding results of very great precision.

All the three methods last mentioned are quite accurate and exhibit an excellent agreement. It is claimed that the charge e on an electron is known to within 1 per cent. Millikan's|| latest value is $e = 4{\cdot}81 \times 10^{-10}$ E.S. unit or $1{\cdot}60 \times 10^{-20}$ E.M. unit.

---

* See chap. XV.

† *Roy. Soc. Proc.* A. vol. LXXXI. pp. 141, 163 (1908).

‡ *Phil. Mag.* VI. vol. v. p. 429 (1903).

§ *Phil. Mag.* VI. vol. XIX. p. 209 (1910).

|| *Phys. Rev.* vol. XXXIV. p. 399 (1912).

Electromagnetic Inertia.

About thirty years ago J. J. Thomson* pointed out the extremely important result that an electric charge possessed inertia simply in virtue of the energy of its electromagnetic field, and he succeeded in calculating the magnitude of the electrical inertia, or electromagnetic mass, as it is usually called, of a charged sphere. This additional mass does not play any important part in ordinary electrostatic experiments, as it is always small compared with the mass of the uncharged portion of the conductors and insulators which are experimented with. The case is very different when we are dealing with charged particles whose mass is only one eighteen-hundredth of that of an atom of hydrogen. It was obviously an important experimental problem to determine how much of the inertia of the electron was of the type foreseen by Thomson and how much, if any, was to be attributed to "ordinary" mechanical mass. Fortunately the two kinds of mass differentiate themselves rather clearly. The mechanical mass is supposed to be independent of the velocity of the body, following the principles of mechanics laid down by Newton, whereas the electromagnetic mass is a continuous function of the velocity and approaches infinity as the velocity of the electric charge approaches that of light.

The experimental problem was resolved by Kaufmann†, who measured the value of e/m for the electrons emitted by radium bromide, some of which have velocities as high as $2{\cdot}89 \times 10^{10}$ centimetres per second. He showed that the mass of these electrons varied with the speed, and in fact in a manner very similar to that predicted by Thomson. His final conclusion was that if there was any part of the mass of an electron which was ordinary mechanical mass it was very small in comparison with the part which was of electromagnetic origin.

The Negative Electron.

We have now succeeded in the isolation of a charged particle whose mass is much less than that of any known chemical atom. So far as our experiments enable us to tell, the whole of this mass arises from the electric charge the particle carries. There can be no question but that this is the negative electron. The structure of the positive electricity which goes to make up the remainder of the uncharged chemical atom is still uncertain; but the results of experiments so far point to both the charge and mass of the positive electrons being different from those of the negative. There is no evidence, so far as the writer is able to observe, which supports the view that the positive electron is, as it were, a reflection of the negative.

---

* *Phil. Mag.* V. vol. XI. p. 229 (1881).

† *Ann. der Phys.* vol. XIX. p. 487 (1906); cf. also H. Starke, *Ver. der Deutsch. Physik. Ges.* vol. v. p. 241 (1903).

The foregoing considerations enable us to define more precisely the use of the word electron. In future we shall restrict the term to particles which consist of a geometrical configuration of electricity and nothing else, whose mass, that is, is all electromagnetic. For a particle which is a charged molecule or atom, that is to say, a molecule or atom which has lost or gained one or more negative electrons, we shall use the term ion. A wider meaning than this is currently attributed to the word ion, in the sense of any charged particle which is considered to have a separate existence or which behaves as a dynamical unit. Some of these are comparatively large bodies and contain very many atoms or molecules. To distinguish them from the smaller ions already referred to, Stark has suggested the use of the terms molion and atomion. As, however, we shall not have to consider the large molions we shall simply use the word ion instead of atomion.

According to the view we are developing, all interactions between material systems result from the electrical charges which make up their ultimate parts. The space in the neighbourhood of an electric charge is to be looked upon as having properties different from that some distance away, since an electric charge of the same sign is repelled with a greater force in the former case than in the latter. This state of things is described by saying that the electric charge is surrounded by a field of force.

It is often convenient to attribute this field of force to disturbances produced by the electric charge in a medium, the aether, which fills all space. Looked at in this way the real electron, the part which acts, is the surrounding aether which is outside its geometrical boundary; and the electron theory is the science of the properties of the aether, of which the electric charges are local modifications.

Different Elements of Electricity.

Our ignorance of the geometrical distribution of the electrification constituting an electron is almost complete, but this is not a serious disadvantage in considering many applications of the theory. It will often be sufficient to regard an electron as a point charge having a definite inertia coefficient or mass. In such cases the mode of distribution of the electric charge, whether it is a point, line, surface or volume charge and whether it is distributed with spherical or linear symmetry or not, is unimportant, provided that it is confined to a minute region of space. Although this is often true there are some investigations for which the ultimate geometrical distribution of the electrification is important; as for instance in the case already mentioned of the calculation of the electromagnetic mass. These two distinct classes of cases require as a rule quite different methods of attack.

These remarks will make it clear why it is necessary to have different elementary portions of electricity in different investigations dealing with the electron theory. In the first place we may have to consider the forces acting on or arising from a small portion of the electron itself. The elementary quantity of electricity concerned here may be denoted by $\rho\, d\tau$, where ρ is the volume density of the electricity at the point of the electron under consideration and $d\tau$ is an element of volume of the latter, infinitesimal in comparison with the size of the electron. In other investigations our element of electricity will be the charge on a single electron whose value e is determined by the equation

$$e = \iiint \rho\, d\tau,$$

where the volume integral extends over the volume occupied by the electron. In still another class of investigations the volume element of electric charge will occupy a region of space containing an enormous number of electrons both negative and positive. In this case also it is conveniently represented by $\rho\, d\tau$, or when confusion is likely to occur with the first case by $\bar{\rho}\, d\tau$. If n is the number of negative electrons per unit volume at any point and e is the charge carried by each, N and E being the corresponding quantities for the positive electrons, then $\bar{\rho}\, d\tau = (NE + ne)\, d\tau$.

The function E is what is known as a vector point function; in other words it is a function which for each point of space has both magnitude and direction. A function which has only one magnitude at each point may be called a single-valued point function. The electric intensity E is single-valued in this sense; but its direction is indefinite at the points where E vanishes.

Clearly the electric intensity at a point distant r from a point charge e' is in Heaviside's units $E = e'/4\pi r^2$. This follows from the inverse square law of force. The electric intensity due to a complicated distribution of charge may obviously be obtained by the integration of the amount arising from each volume element. In making the calculation it is necessary to integrate for each component of the electric intensity separately and combine the results according to the rule for the composition of forces. This resolution and subsequent composition of vectors is often troublesome as well as clumsy, and it is not necessary for the calculation of the electric intensity. It may be dispensed with by the introduction of another function known as the Potential.

The Potential.

The Potential is defined as the work divided by the charge when an infinitesimal electric charge is brought from some standard position to the point in question. The standard position is usually taken to be a point at an infinite distance away from charged bodies. The value of the potential calculated in this way must be independent of the path of approach to the point under consideration, otherwise an indefinite amount of work could be obtained by making the charge move round a closed contour passing through the point under consideration and the standard position. This would be contrary to the law of conservation of energy. The potential is a function of each point in space and possesses magnitude but not direction. Such a function is known as a *scalar* point function. The electrostatic potential is single-valued.

Let P, Q be two points at an infinitesimal distance ds apart and such that PQ is in the direction of the resultant electric intensity E at P. Let dV be the increase in the potential in passing from P to Q and let the direction $P \rightarrow Q$ be considered positive. Then add to the potential an amount $\frac{1}{4\pi}\iint \frac{\sigma}{r}\, dS$, where σ is the surface density of the charge at the element dS and r its distance from the external point.

The Potential at internal points.

Further consideration is necessary before we can apply the foregoing expression to the potential at points inside the electrified medium. At such points the denominator r becomes zero, and we must be certain that this does not make the integral which represents the potential infinite. The matter may be investigated as follows. About the internal point Q describe a sphere of infinitesimal radius ϵ. We shall suppose the density ρ of the charge to be everywhere finite and the radius ϵ to be chosen so small that the density of the charge varies continuously throughout the volume of the sphere. This condition can always be satisfied. The potential at Q will consist of two parts: (1) V_1, arising from the charge outside the sphere, and (2) V_2, arising from the charge inside the sphere. The former is clearly finite. Let ρ_m be the maximum value of ρ inside the sphere. Then $V_2 \ngtr \frac{1}{4\pi}\iiint_\epsilon \frac{\rho_m}{r}\, d\tau$. The element of volume $d\tau$ in polar coordinates is $r^2 dr \sin\theta\, d\theta\, d\phi$. So that

$$V_2 \ngtr \frac{\rho_m}{4\pi}\int_0^\epsilon r\, dr \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\phi \ngtr \tfrac{1}{2}\rho_m \epsilon^2.$$

This vanishes when ϵ is made sufficiently small; so that we conclude that the charge in the immediate neighbourhood contributes nothing to the potential at any point. The formula $V = \frac{1}{4\pi}\iiint \frac{\rho}{r}\, d\tau$ therefore holds generally both for points outside and for points inside the electrified medium.

The Derivatives of V.

The electric intensity $-\frac{\partial V}{\partial r}$ and all higher derivatives of V contain r to a higher order in the denominator than does V itself. They are therefore all finite at external points.

The electric intensity is also finite at internal points. Consider again the small sphere of radius ϵ and divide V into two parts

Consider first the intensity E_1 at Q due to a charge e_1 at a point P within the surface. We have $E_1 = e_1/4\pi PQ^2$ and

$$N_1 dS = \frac{e_1}{4\pi PQ^2} dS \cos\theta,$$

where θ is the angle between PQ and the normal to dS. Now $\frac{dS\cos\theta}{PQ^2}$ is the solid angle $d\omega_1$ subtended by dS at P. Thus, so far as the single charge e_1 is concerned, we have $N_1 dS = \frac{e_1}{4\pi} d\omega_1$. If there are a large number of point charges e_1, e_2, etc., the resultant normal component of electric intensity

$$N = N_1 + N_2 + N_3 + \dots,$$

so that $$4\pi N\,dS = e_1 d\omega_1 + e_2 d\omega_2 + e_3 d\omega_3 + \dots.$$

Now $$\iint e_1 d\omega_1 = 4\pi e_1, \quad \iint e_2 d\omega_2 = 4\pi e_2, \text{ etc.}$$

So that $$\iint N dS = e_1 + e_2 + e_3 + \dots = e \quad \dots\dots\dots\dots\dots(3).$$

This result can obviously be extended from a series of point charges to a continuous distribution in the same manner as that employed in dealing with the potential.

It remains to prove that charges outside the closed surface contribute nothing to the surface integral. It is evident that every conical element of solid angle $d\omega$ arising from an external charge will cut the closed surface an even number of times. The value of NdS for the intersections of the cone by the surface will be alternately positive and negative since the direction of the electric intensity is constant in space but alternates in sign with reference to the successive normals. The numerical value of NdS is the same for successive intersections, being equal to $\frac{e}{4\pi} d\omega$. So that the surface integral is divided up into a series of pairs of equal and opposite elements. Its value is therefore zero, and we conclude that the value of $\iint NdS$ over any closed surface is equal to the charge inside.

This result is known as Gauss's Theorem.

The theorem is of great value as a means of calculating the value of the electric intensity arising from various symmetrical distributions of electric charge. Thus in the case of a uniformly charged spherical shell the intensity at any point external to the shell must be the same at every point on the spherical surface through the point and concentric with the shell. It also follows from the symmetry of the problem that the intensity must be radial. Since the area of a sphere of radius r is $4\pi r^2$, the electric intensity E at a distance r from the centre of the charged shell is given by $4\pi r^2 E = e$; so that

$$E = e/4\pi r^2 \quad\text{.........................(4).}$$

In a similar way we may prove that the force vanishes inside the shell. These results may be extended to the case of a sphere charged throughout its volume so that layers equidistant from the centre are charged to equal density. Thus we may show, for example, that the force *inside* a solid uniform sphere of electricity varies as the distance from the centre.

The application of Gauss's Theorem to the tubes of force mentioned on p. 16, is instructive. As we have seen, a tube of force is a tubular region bounded by a surface which is the envelope of the lines of force. Let us apply Gauss's Theorem to a portion of such a tube, terminated at each end by equipotential surfaces. The lines of force run along the tubular surfaces so that at each point the component of the intensity normal to these surfaces vanishes. Over the ends the resultant electric intensity will be normal to the surfaces. Let it be E_1 at the end where the cross section is S_1, E_2 and S_2 being the corresponding quantities at the other end. The value of $\iint N dS$ over the whole surface considered is clearly $E_1 S_1 - E_2 S_2$. If the tube of force is in a region where there are no electric charges this vanishes, so that $E_1 S_1 = E_2 S_2$; thus the electric intensity at any point is inversely as the area of cross section of the tubes of force at that point.

Under the conditions contemplated in electrostatics the surface of a conductor of electricity must be an equipotential surface; otherwise there would be currents of electricity flowing from one part of the surface to another. The tubes of force must therefore start normally from such a surface. Now apply Gauss's Theorem to the region bounded by a tube of force and its continuation into the substance of the conductor and terminated by equipotential surfaces, one inside and the other outside the conductor. The electric intensity vanishes over all the surface inside the conductor and the normal component vanishes over the tubular surface outside. The value of $\iint N dS$ is thus equal to the value $E_1 S_1$ of this quantity over the end section. This is equal to the charge inside; which is σS_2, where σ is the surface density of the charge on the conductor and S_2 the area of its intersection by the tube of force. We therefore have $E_1 S_1 = \sigma S_2$. If we make the end section approach indefinitely near to the charged surface, $S_1 = S_2$, so that

$$E = \sigma \quad \text{.............................} \quad (5).$$

Thus the electric intensity at a point just outside a charged conductor is normal to the surface and equal to the charge per unit area of it. This result is known as Coulomb's Law.

We may express Gauss's Theorem analytically as a relation between a surface and a volume integral. If ρ is the volume density of the electrification at any point inside a closed surface the total charge inside the surface will be $\iiint \rho\, dx\, dy\, dz$. The normal component of the electric intensity outside the surface is $N = -\frac{\partial V}{\partial n}$. Gauss's Theorem may therefore be expressed in the form

$$\iiint \rho\, dx\, dy\, dz = -\iint \frac{\partial V}{\partial n} dS \quad \text{.................} \quad (6).$$

In a region where there are no charges, if l, m, n are the direction cosines of the normal to any closed surface, we have

$$\iint (lE_x + mE_y + nE_z)\, dS = \iint N\, dS = 0.$$

A vector E whose components E_x, E_y, E_z satisfy the relation $\iint (lE_x + mE_y + nE_z)\, dS = 0$ over any closed surface is said to be solenoidal. Thus the electric intensity in free aether is a solenoidal vector. We shall see that there are other solenoidal vectors in the theory of electricity and magnetism whose properties are of great importance.

Green's Theorem.

An important theorem discovered by George Green* in 1828 enables us to express a volume integral taken throughout an enclosed space in terms of surface integrals over the boundaries of the space. This theorem, which is named after the discoverer, is a purely geometrical theorem, but it has many important applications in the theory of electricity.

---

* **"An essay on the application of Mathematical Analysis to the theories of Electricity and Magnetism," by George Green, published at Nottingham in 1828.**

The total flux is therefore

$$-r^2\,dr\sin\theta\,d\theta\,d\phi\left\{\frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2}+\frac{\cot\theta}{r^2}\frac{\partial V}{\partial\theta}\right\}.$$

The mean potentials of the plane surfaces *PQTU* and *RSVW* are $V+\frac{1}{2}\frac{\partial V}{\partial\phi}d\phi$ and $V-\frac{1}{2}\frac{\partial V}{\partial\phi}d\phi$ respectively. The element of arc normal to the planes being $r\sin\theta\,d\phi$, the mean normal intensities are

$$-\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left(V+\frac{1}{2}\frac{\partial V}{\partial\phi}d\phi\right)\text{ and }+\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left(V-\frac{1}{2}\frac{\partial V}{\partial\phi}d\phi\right).$$

The areas of the surfaces are equal to each other and to $dr\times r\,d\theta$. The fluxes are therefore equal to

$$-\frac{dr\,d\theta}{\sin\theta}\left(\frac{\partial V}{\partial\phi}+\frac{1}{2}\frac{\partial^2 V}{\partial\phi^2}d\phi\right)$$

and

$$+\frac{dr\,d\theta}{\sin\theta}\left(\frac{\partial V}{\partial\phi}-\frac{1}{2}\frac{\partial^2 V}{\partial\phi^2}d\phi\right),$$

their sum being

$$-r^2 dr\sin\theta\,d\theta\,d\phi\times\frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2}.$$

We thus find for the total flux over all the six surfaces of the element of volume

$$-r^2 dr\sin\theta d\theta d\phi\times\frac{\partial^2 V}{\partial r^2}+\frac{2}{r}\frac{\partial V}{\partial r}+\frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2}+\frac{\cot\theta}{r^2}\frac{\partial V}{\partial\theta}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2}.$$

By Gauss's Theorem this is equal to the charge inside, which is

$$\rho\,d\tau=\rho\times r^2\sin\theta\,dr\,d\theta\,d\phi.$$

So that

$$\frac{\partial^2 V}{\partial r^2}+\frac{2}{r}\frac{\partial V}{\partial r}+\frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2}+\frac{\cot\theta}{r^2}\frac{\partial V}{\partial\theta}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2}=-\rho\ldots(14).$$

This is therefore Poisson's Equation in spherical coordinates, and the operator

$$\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{\cot\theta}{r^2}\frac{\partial}{\partial\theta}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\quad\ldots\ldots(15)$$

is the form which Laplace's operator ∇^2 takes in this system of coordinates.

Cylindrical Coordinates.

Any other system of orthogonal coordinates may be treated similarly. For instance in the case of cylindrical coordinates r, θ, z the element of volume is bounded by (1) two coaxal cylinders of radii r and $r+dr$ with their axes coincident with the axis of z, (2) two planes inclined at an angle $d\theta$ to one another and passing through the axis of z, (3) two parallel planes perpendicular to the axis of z and at a distance dz apart.

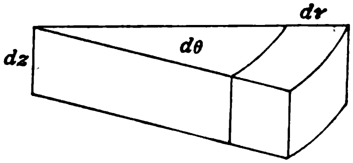


Fig 6.

The volume of the element is clearly $dr \times rd\theta \times dz$. The potential at the centre of the element being V, the flux over the outer cylindrical surface will be

$$\left(r+\frac{dr}{2}\right) d\theta dz \times -\frac{\partial}{\partial r}\left(V+\frac{1}{2}\frac{\partial V}{\partial r}dr\right),$$

and over the opposite face

$$\left(r-\frac{dr}{2}\right) d\theta dz \times \frac{\partial}{\partial r}\left(V-\frac{1}{2}\frac{\partial V}{\partial r}dr\right),$$

the total flux over the two faces being

$$-r dr d\theta dz \left\{\frac{\partial^2 V}{\partial r^2}+\frac{1}{r}\frac{\partial V}{\partial r}\right\}.$$

The area of the plane inclined faces is $dr\,dz$, the mean normal intensity over them $\mp\frac{1}{r}\frac{\partial}{\partial\theta}\left(V\pm\frac{\partial V}{\partial\theta}d\theta\right)$ and the total flux over them

$$-r\,dr\,d\theta\,dz\left\{\frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2}\right\}.$$

The area of the faces perpendicular to Oz is $dr \times rd\theta$ and the total flux over them $= -r\,dr\,d\theta\,dz\,\frac{\partial^2 V}{\partial z^2}$.

So that the flux over the whole six faces is

$$-r dr d\theta dz \left\{\frac{\partial^2 V}{\partial r^2}+\frac{1}{r}\frac{\partial V}{\partial r}+\frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2}+\frac{\partial^2 V}{\partial z^2}\right\} = \rho \times r dr d\theta dz.$$

Thus the form which Poisson's Equation takes in cylindrical coordinates is

$$\frac{\partial^2 V}{\partial r^2}+\frac{1}{r}\frac{\partial V}{\partial r}+\frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2}+\frac{\partial^2 V}{\partial z^2} = -\rho \quad \ldots\ldots\ldots\ldots(16).$$

Since $\iint (V-V')\frac{\partial}{\partial n}(V-V')\,dS$ also vanishes when $\frac{\partial V}{\partial n}$ is given over the surface S, it follows that V is unique except for an additive constant when the value of ρ is assigned throughout the space and that of $\frac{\partial V}{\partial n}$ over the boundaries; so that, in this case also, the electric intensity is determined uniquely.

It would lead us too far afield to consider the functions (Fourier's Series, Spherical and Zonal Harmonics and Bessel's Functions) which are the solutions of Laplace's Equation appropriate to particular problems. For the development of this interesting subject the reader may be referred to Byerly's *Fourier's Series and Spherical Harmonics.*

Total Energy of a System of Charges.

We may find the total energy of a system of charged bodies in terms of their charges and potentials as follows. Since the potential at any point of the field is equal to $\iiint \frac{\rho}{r}\,d\tau + \iint \frac{\sigma}{r}\,dS$ taken over all the charged bodies in the field, it will be reduced to $1/n$ of its value if *all* the charges are reduced in the ratio $n:1$. Let n be any very large number, and suppose that initially all the charges are at an infinite distance from one another. Bring up $1/n$ of each element of charge to its final position. If V is the final potential at any point the potential will change during this operation from 0 to $\frac{V}{n}$. The work done in bringing up the element $\frac{1}{n}\rho\,d\tau$ will lie between 0 and $\frac{1}{n^2}V\rho\,d\tau$. The work done in bringing up $1/n$ of all the charges will lie between 0 and $\frac{1}{n^2}\iiint V\rho\,d\tau$. Now bring up a second nth part of all the charges. This will raise the potential at any point from V/n to $2V/n$, and the work done in this second stage will lie between $\frac{1}{n^2}\iiint V\rho\,d\tau$ and $\frac{2}{n^2}\iiint V\rho\,d\tau$. If this process is continued the work done in the sth stage will lie between

$$\frac{s-1}{n^2}\iiint V\rho\,d\tau \text{ and } \frac{s}{n^2}\iiint V\rho\,d\tau.$$

The total work done in bringing up the whole of the charges from a state of infinite dissemination will lie between

$$\frac{1}{n^2}(0+1+2+\dots+\overline{n-1})\iiint V\rho\,d\tau$$

and
$$\frac{1}{n^2}(1+2+3+\dots+n)\iiint V\rho\,d\tau.$$

This is equal to the total potential energy of the system, which therefore lies between

$$\frac{n(n-1)}{2n^2}\iiint V\rho\,d\tau \text{ and } \frac{n(n+1)}{2n^2}\iiint V\rho\,d\tau.$$

When n is increased indefinitely each of these values coincides with $\frac{1}{2}\iiint V\rho\,d\tau$. This is the part due to the volume charges. If there are surface charges we shall have to add $\frac{1}{2}\iint\sigma V\,dS$. The complete expression for the total energy of any system of charges is therefore

$$W=\frac{1}{2}\iiint V\rho\,d\tau+\frac{1}{2}\iint V\sigma\,dS \quad\dots\dots\dots\dots(17).$$

The Energy in the Field.

In the preceding paragraph we have deduced an expression for the energy in terms of the charges and their potentials. On the view that electrical actions are transmitted through a medium, we should expect that the energy would reside in the medium. It is easy to obtain from the equation (17) an expression which admits of this interpretation. Since $\rho=-\nabla^2V$ and $\sigma=-\frac{\partial V}{\partial n_e}$, we have

$$W=-\frac{1}{2}\iiint V\nabla^2V\,d\tau-\frac{1}{2}\iint V\frac{\partial V}{\partial n_e}\,dS.$$

But by Green's Theorem, allowing for the reversal of sign arising from the fact that the normal is now drawn *into* the space considered, this is equal to

$$W=\frac{1}{2}\iiint\left\{\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2\right\}d\tau \quad\dots\dots(18)$$

$$=\frac{1}{2}\iiint E^2\,d\tau \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(19).$$

So that the energy of the system is the same as if each element of the field contained an amount of energy $\frac{1}{2}E^2$ per unit volume.

Thus the integrand will be in the form desired if we put

$$P = \frac{1}{2}\left\{\left(\frac{\partial V}{\partial x}\right)^2 - \left(\frac{\partial V}{\partial y}\right)^2 - \left(\frac{\partial V}{\partial z}\right)^2\right\} = p_{xx} \quad \ldots\ldots(20),$$

$$Q = \frac{\partial V}{\partial x}\frac{\partial V}{\partial y} = p_{yx} = p_{xy} \quad \ldots\ldots(21),$$

$$R = \frac{\partial V}{\partial x}\frac{\partial V}{\partial z} = p_{zx} = p_{xz} \quad \ldots\ldots(22).$$

By considering the components of the resultant force parallel to the y and z axes we should arrive at similar surface integrals involving

$$p_{yy} = \frac{1}{2}\left\{\left(\frac{\partial V}{\partial y}\right)^2 - \left(\frac{\partial V}{\partial z}\right)^2 - \left(\frac{\partial V}{\partial x}\right)^2\right\} \quad \ldots\ldots(23),$$

$$p_{zz} = \frac{1}{2}\left\{\left(\frac{\partial V}{\partial z}\right)^2 - \left(\frac{\partial V}{\partial x}\right)^2 - \left(\frac{\partial V}{\partial y}\right)^2\right\} \quad \ldots\ldots(24),$$

$$p_{zy} = \frac{\partial V}{\partial y}\frac{\partial V}{\partial z} = p_{yz} \quad \ldots\ldots(25),$$

p_{xy} and p_{xz}.

In the new notation in terms of the p's, which will be familiar to students of elasticity, we may write

$$+X_1 = \iiint\left(\frac{\partial p_{xx}}{\partial x} + \frac{\partial p_{yx}}{\partial y} + \frac{\partial p_{zx}}{\partial z}\right)d\tau$$

$$= \iint(lp_{xx} + mp_{yx} + np_{zx})\,dS \quad \ldots\ldots(26).$$

The last integral is taken over the enclosing surface, and l, m, n are the direction cosines of dS, drawn away from the enclosed volume. Similarly

$$+Y_1 = \iint(lp_{xy} + mp_{yy} + np_{zy})\,dS \quad \ldots\ldots(27)$$

and

$$+Z_1 = \iint(lp_{xz} + mp_{yz} + np_{zz})\,dS \quad \ldots\ldots(28).$$

If we adopt the standpoint that the action of the electric charges on one another is transmitted by the intervening medium, then $p_{xx}, p_{yy}, p_{zz}, p_{yz}, p_{zx}, p_{zy}$ are the six components of the stress which transmits the action. From the point of view of action at a distance these quantities, on the other hand, have no physical significance.

now have a negative sign in front of them, they will represent a pressure, not a tension. There is thus at every point in the field a tension equal to $\frac{1}{2}E^2$ along the lines of force and a pressure of equal amount in every direction at right angles to them.

It may be shown that these stresses will keep the aether in equilibrium. They are not the most general system of stresses which are equivalent to the electric force on the system of charged bodies e_1. For we may clearly add to them any distribution of stress whose resultant is equal to zero when integrated over any closed surface surrounding e_1. They are, however, the only system in which the stress at any point is determined solely by the electric intensity at that point*.

It is well to point out that this interpretation of the forces as a system of stresses in the medium has only been shown to be a possible, not a necessary one.

---

* Cf. Jeans, *Electricity and Magnetism*, p. 146, and Maxwell, *Electricity and Magnetism*, 3rd ed. vol. I. p. 158.

CHAPTER III

DIELECTRIC MEDIA

CAVENDISH, and Faraday independently, showed that when a condenser was charged so that the potential difference between the plates had a certain fixed value, the charge on the plates depended on the insulating medium between them. This proved that the forces between electric charges depended not only on the magnitude of the charges, but also on the nature of the material separating them. These experiments are often regarded as disproving the dogma of action at a distance. They are not capable, however, of establishing this inference; all that they prove with certainty is that electric charges act on a *material* medium in such a way as to make it affect other electric charges. The quantitative experiments of the investigators mentioned showed that the charge on condensers of different geometrical form to which a given difference of potential was applied always changed in a certain ratio when any assigned insulating material was replaced by any other assigned insulating material. Different insulators were therefore said to be characterized by different "specific inductive capacities." The specific inductive capacity of a vacuum is now universally adopted as a standard, and its value is put equal to unity on the electrostatic system of units. The specific inductive capacity of air only differs very slightly from it. In this book, for reasons which will appear, we shall use the term dielectric coefficient instead of specific inductive capacity.

The potential difference between two electric charges is, by its definition, determined by the distribution of the electric intensity in the field surrounding them. It is clear, therefore, that these experiments prove that the forces between charged bodies are not determined solely by the magnitude of their charges and their distance apart. Those of our previous results the same as that of those which cross an area $\cos\theta$ perpendicular to OP, so that

$$D_\theta = D\cos\theta.$$

It is evident that the cross-sectional area of the tube of induction passing through any point is the inverse of the induction at that point.

Gauss's Theorem.

We are now in a position to consider the form which Gauss's Theorem takes when the field embraces dielectric media. Consider the value of the integral $\iint D_n dS$ taken over any closed surface, where D_n is the component of the induction along the outward

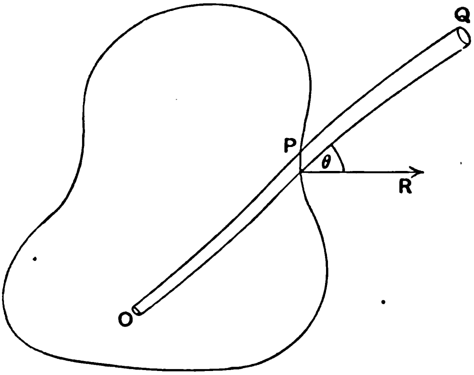


Fig. 8.

drawn normal at an element dS of the surface. If OPQ represents the direction of the tubes crossing dS, and if the angle $RPQ = \theta$, then $D_n = D\cos\theta$, where D is the resultant induction at dS. Now $dS\cos\theta$ is the projection of dS on a plane perpendicular to the direction of D, so that

$$D \times dS\cos\theta = D_n dS$$

is the number of tubes which cross dS from the inside to the outside of the surface. When the tubes cross from the outside to the inside of the closed surface $D_n dS$ will have a negative value. Let us first consider the effect of those elements for which $D_n dS$ has a positive value. Each of the tubes crossing them will start from a unit positive charge beyond dS on the internal side. This unit charge may either be without or within the surface. If it is without, the tube will first cross the surface at some other point, and it will give rise to an equal and opposite element $D_n'dS'$. The total contribution of such tubes to the integral will be zero. Each tube crossing an element dS in the positive direction and having its origin within the surface will start on a unit positive charge, so that the total contribution of all the elements $D_n dS$ which are positive will be equal to that part of the positive charge within the surface which gives rise to tubes of force which do not end on a negative charge inside. In a precisely similar manner we can show that the total contribution of all the elements $D_n dS$ which are negative is equal to that part of the negative charge within the surface on which tubes of force end which do not start from positive charges inside the surface.

It follows that the value of $\iint D_n dS = \iint D\cos\theta\, dS$ taken over any closed surface is equal to the algebraic sum e of all the charges enclosed by the surface.

We thus see that Gauss's Theorem can be extended to dielectric media provided the normal electric intensity is replaced by the normal induction.

By applying Gauss's Theorem to a cylindrical region bounded by a tube of force and by two equipotential surfaces, one just inside and the other just outside a conductor, we find that the induction just outside the surface of a conductor is along the normal to the surface and equal to the charge per unit area of it. This is the general form of Coulomb's Law.

Poisson's and Laplace's Equations.

Let us now apply Gauss's Theorem to the rectangular parallelepiped whose angular points are the combinations of $x \pm \frac{1}{2}dx$, $y \pm \frac{1}{2}dy$, $z \pm \frac{1}{2}dz$. The induction at the centre of the element being $D\,(D_x, D_y, D_z)$, the mean outward normal induction over the faces perpendicular to the axis of x will be $-\left(D_x - \frac{\partial D_x}{\partial x}\frac{dx}{2}\right)$ and $+\left(D_x + \frac{\partial D_x}{\partial x}\frac{dx}{2}\right)$ by Taylor's Theorem. The corresponding fluxes of induction are

$$-\left(D_x - \frac{\partial D_x}{\partial x}\frac{dx}{2}\right)dy\,dz$$

and

$$+\left(D_x + \frac{\partial D_x}{\partial x}\frac{dx}{2}\right)dy\,dz.$$

two sides of it, by an amount equal to the charge per unit area of the surface at that point. In addition, the tangential component of the electric intensity must have the same value on both sides of the surface. There is only one solution of the differential equation which satisfies the conditions in any assigned case; so that any particular solution of the differential equation which can be made to satisfy the boundary conditions will determine the field uniquely. The boundary conditions which we have established for dielectric media include those relating to conductors in a vacuum, or in any dielectric, as particular cases.

Poisson's Theory of Dielectric Media.

There is a certain view of the behaviour of dielectric media the mathematical development of which is largely due to the French physicist Poisson. According to this view the modification of the electric field arising from the presence of dielectrics is due to the substance of the dielectric being thrown into a peculiar electrical condition by the external field. This condition arises from the displacement of electricity in the ultimate particles of the medium in such a way that each particle acquires a positive charge at one end and a negative charge at the other. When this occurs the medium is said to be polarized and we shall see that the polarization is measured by the product of the displacement and the charge per unit volume.

We shall first of all consider the nature of polarization as it presents itself from the point of view which regards electricity as continuously distributed, after the manner of a fluid, throughout all bodies. Consider two equal spheres, and let one of them be filled with a uniform distribution of positive electricity and the other with a distribution of negative electricity, identical with the first except for the difference of sign. Imagine the two spheres coincident in position; we shall then have an uncharged body which will give rise to no electrical effects. Now suppose that, whilst one of the spheres is fixed, every point of the other is given a certain equal displacement, so that this sphere moves a small distance after the manner of a rigid body. The region of space where the spheres overlap will still be free from electric charge, but there will be a layer of positive electricity over the surface of the sphere on that side towards which the positive electricity has been

A body constructed in this way is identical electrically with a body made up of two uniform distributions of positive and negative electricity which have undergone relative displacement. This is clear, because in the interior of the body equal and opposite charges coincide at the contiguous surfaces of opposite elements, so that there is no volume charge at any point; whilst at the external boundary there is a charge, positive at one end of the body and negative at the other end, whose magnitude is the same for each portion of the surface in which it is intersected by the sides of a cell. But the area of this intersection is inversely as the cosine of the angle that the normal to the surface makes with the axis of the cells. Thus the density of this fictitious surface layer varies as the cosine of the angle that the normal makes with the axis of polarization, and the distribution is completely identical with that obtained by displacing originally coincident electric charges.

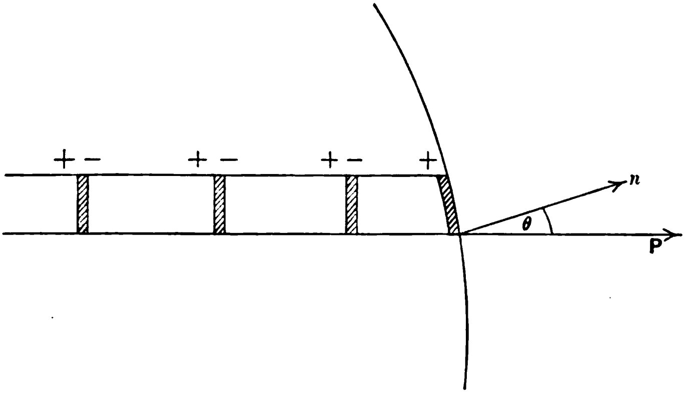


Fig. 10.

In some problems it is more convenient to regard polarized media as displaced charges which were originally coincident, whilst in others the point of view which considers them as made up of cellular elements having opposite charges on opposite faces has advantages.

If we suppose a polarized body to be intersected by any surface, the two resulting portions of the body will still be polarized. It is necessary therefore that there should be developed at any such surface of separation a double distribution of electric charge, of equal amount and opposite sign at any point of the interface.

of electric charge is $\cos\theta\,\delta x$, if δx is the length of a cell. The electric moment (perhaps strength would be a better term) of the slab is therefore $\sigma\cos^2\theta\,\Delta S\,\delta x$. The volume of the slab is

$$\cos\theta\,\Delta S\,\delta x,$$

so that the polarization along the direction normal to the slab is clearly

$$P_\theta = \sigma\cos\theta = P\cos\theta \quad \text{..................(12)},$$

where P is the resultant polarization. The components of the polarization are therefore obtained by the usual rule for the resolution of vectors.

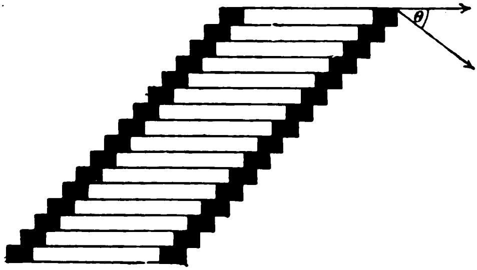


Fig. 11.

Polarization, Induction and Electric Intensity.

We are now in a position to determine the relation between these three vectors. Let us first consider the nature of the electric field between the plates of a parallel plane condenser filled with a dielectric of specific inductive capacity κ. Then the induction D and the electric intensity E are both normal to the plates and $\kappa E = D = \sigma$, where σ is the charge per unit area of the plates.

We shall make the hypothesis that the polarization is caused by the electric intensity in the dielectric, and that the two vectors are coincident in direction. This standpoint will be found to be fully justified when we come to consider the phenomena from the point of view of the electron theory, according to which the polarization arises from an actual displacement of the ultimate electrified particles in the direction of the field. The hypothesis is undoubtedly true for isotropic dielectric substances. In the case of crystalline substances the electric intensity does not in charge due to the polarization and equal to $\pm P$ per unit area. The force on a unit charge in such a cavity will arise partly from the charge on the walls and partly from the charges, real and fictitious, in the rest of the field. The former part is equal to P and the latter to E. The total force is therefore $P+E=D$. The resultant induction D at any point is therefore equal to the force which would act on a unit positive charge placed in a very flat cavity cut perpendicular to the direction of the lines of force at that point.

It is easy to show that the component of the induction in any direction is the component, in that direction, of the force which would be exerted on a unit positive charge placed in a similar flat cavity with its end faces normal to the direction in question.

The foregoing specifications of the induction and electric intensity satisfy the conditions which we have already laid down for them (p. 45) at the interface between two media K_1 and K_2 of different dielectric coefficients. For, consider two flat cavities A and B parallel to the interface C and indefinitely near to it. The normal component of the induction just inside K_1 will be equal to the force on a unit charge in A and that just inside K_2 will be equal to the force acting on a similar charge in B. If the resultant intensities in the two media are E_1 and E_2, and the resultant polarizations are P_1 and P_2, and if they make angles θ_1 and θ_2 with the normals to the surface, then it follows from the results of the preceding paragraph that the normal force in A is $(E_1+P_1)\cos\theta_1$, and that in B is $(E_2+P_2)\cos\theta_2$. That these forces are equal to one

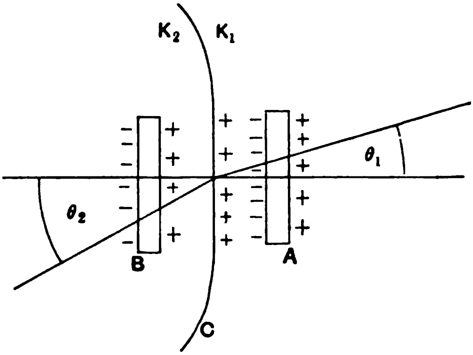


Fig. 12.

charges, ρ' and σ' of the polarized distribution. The electric intensity will therefore obey the form of Poisson's equation which obtains in the free aether, provided the density at any point is supposed to be equal to $\rho + \rho'$. We therefore have

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \rho + \rho' \quad \text{.................(18).}$$

Now
$$\rho' = -\left(\frac{\partial P_x}{\partial x} + \frac{\partial P_y}{\partial y} + \frac{\partial P_z}{\partial z}\right),$$

so that
$$\frac{\partial}{\partial x}(E_x + P_x) + \frac{\partial}{\partial y}(E_y + P_y) + \frac{\partial}{\partial z}(E_z + P_z) = \rho \quad \text{....(19).}$$

Comparing this with equation (1), we see that the vector whose components are $E_x + P_x$, $E_y + P_y$, $E_z + P_z$ is identical with the induction D, so that the identification from a particular case on p. 52 is perfectly general.

By integrating both sides of equation (19) throughout any enclosed volume we see that the true charge inside is equal to $\iint D_n dS$ over the bounding surface, in agreement with p. 42.

The Fictitious Charges on the Electron Theory.

Although it is necessary, in discussing the results of electrostatic experiments, to distinguish between "true" charges like those which are communicated from a conductor to the plates of a condenser and the "fictitious" charges which appear to reside in the dielectric, there is no very profound difference between them. According to the electron theory one is just as true a charge as the other, although its reality is not so readily made obvious by experiment. The electron theory regards a dielectric as a certain type of distribution of electrons in space, and in this space the true electric intensity satisfies the equation div $E = \rho$. This equation is assumed to be true when the element of volume is a small part of an electron. When the element of volume is enlarged so as to contain a great many electrons the equation will become div $\bar{E} = \bar{\rho}$, where the bars denote average values. Thus $\bar{\rho}$ is equal to the $\rho + \rho'$ of equation (18) and ρ' is just as real a part of the average density of electrification as ρ. This point will be considered more fully in the sequel.

the force exerted on a unit charge occupying an infinitesimal volume, will constantly change in both magnitude and direction from one part of this space to another. At places which are sufficiently close to an electron the actual force will be enormous compared with what we have called the electric intensity, and it may have any direction whatever.

One is tempted to ask what can be the use of a conception of the electric intensity which is so much at variance with what we believe to be the reality. The answer is, of course, that most of our methods of experimenting are so coarse, compared with the atomic scale, that they do not detect these enormous differences which occur within distances of the order of atomic magnitudes. Our experimental arrangements for the most part measure only the average values over spaces containing a large number of atoms. The reason why our average values possess validity is not because they are the true values but because, so far as such experimental arrangements enable us to detect, everything happens as if the average values were the true values.

It remains to specify the average values we have been dealing with more accurately than we have done hitherto. Let ϕ represent one of the scalar functions or a component of one of the vectors, which determine the state of the electric field. For example, ϕ may be the electrostatic potential at a point. Let τ be any small volume so chosen that its linear dimensions are large compared with atomic distances but small compared with the distances within which changes in ϕ are perceptible by the usual experimental methods. Then the average value of ϕ may be defined as the value of

$$\bar{\phi} = \frac{1}{\tau}\iiint \phi \, d\tau \quad \text{.........................(12)},$$

where the integral is taken throughout the small volume τ. We evidently have

$$\overline{\frac{\partial \phi}{\partial \xi}} = \frac{1}{\tau}\iiint \frac{\partial \phi}{\partial \xi} d\tau = \frac{1}{\tau}\frac{\partial}{\partial \xi}\iiint \phi \, d\tau = \frac{\partial \bar{\phi}}{\partial \xi} \quad \text{.........(13)},$$

where ξ is any independent variable such as time, distance, etc. of which ϕ may be a function.

Since the actual potential V satisfies the relation $\nabla^2 V = -\rho$, it follows that $\nabla^2 \bar{V} = -\bar{\rho}$; and since $E = -\text{grad } V$, $\bar{E} = -\text{grad } \bar{V}$. Thus the average forces and potentials are the same as those which would obtain if the actual charges were replaced by a distribution of density equal to the average density at every point.

It is clear that the induction, polarization and electric intensity in a dielectric are average values in the sense indicated, and that the results that we have deduced are valid if this is understood.

Potential due to the Displaced Electrons.

We have seen that in the presence of an electric field the electrons are displaced, the positive in the direction of the field and the negative in the opposite direction. We shall see that the displacement thus produced is equivalent, for each electron, to the creation of a doublet of moment $\mu = es$, where e is the charge and s the displacement, of that electron. This doublet will contribute to the potential at a distant point P an amount $\frac{\mu}{4\pi}\frac{\partial}{\partial s}\left(\frac{1}{r}\right)$, and if there are ν such doublets per unit volume the total potential to which they will give rise at the point P will be

$$\frac{1}{4\pi}\iiint \nu\mu \frac{\partial}{\partial s}\left(\frac{1}{r}\right) d\tau.$$

In general the different electrons in the atom will be variously situated so that they will not all undergo the same displacement s in a given electric field. We may divide them up into classes, all the electrons in a class being characterised by a given value of s for a given field. Suppose there are n such classes and let ν_p, μ_p and s_p denote the values of ν, μ, s for the electrons of the pth class. Then it is clear that

$$4\pi V_P = \sum_{p=1}^{p=n} \iiint \nu_p \mu_p \frac{\partial}{\partial s_p}\left(\frac{1}{r}\right) d\tau \quad \ldots\ldots\ldots\ldots\ldots(14).$$

We shall now consider the relation between the moment μ_p of the doublets and the electric field which produces them. The exact form of the relation between the restoring force and the displacement will depend on the arrangement of the electrons in the atom. At present our knowledge of this arrangement is very limited but, in any event, the restoring force must be a function of the displacement, which vanishes when the displacement is zero and is opposite in sign to the displacement to which it corresponds.

It follows from Taylor's Theorem that for small displacements the restoring force must be proportional to some odd power of the displacement; and since the frequency of the natural vibrations of bodies, as exemplified by their optical properties, is independent of the amplitude, it is natural to suppose that this power is the first. We shall assume, therefore, that when an electron of the pth class is displaced a small distance s_p the restoring force is equal to $+\frac{1}{\lambda_p} s_p$, where λ_p is a positive quantity which is constant and characteristic of this class of electrons. When the state of equilibrium is attained, the pull of the external electric field on the electron is balanced by the restoring force. The x component of the force on the electron is $-e_p \frac{\partial V}{\partial x_p}$, where V is the actual part of the potential at the electron whose charge is e_p which arises from the presence of the external field. If x_p is the x component of s_p then, provided the reaction to the displacement is independent of its direction in space, the x component of the restoring force measured in the positive direction of x is $-\frac{1}{\lambda_p} x_p$ and the equilibrium value of this displacement is

$$x_p = -\lambda_p e_p \frac{\partial V}{\partial x_p} \quad \text{.....................(15).}$$

Now a moment's consideration shows that when a charge e_p is displaced a distance x_p the electrical effect is exactly the same as that which is produced by the *creation* of a doublet whose moment is $e_p x_p$. For the displaced system is absolutely identical with that which is obtained when a doublet consisting of charges $\pm e_p$ at a distance x_p apart is superposed on the original system, in such a way that x_p coincides with the displacement x_p and the charge $-e_p$ coincides with the original charge $+e_p$. Thus the displacement (15) is equivalent to the creation of a doublet whose moment is

$$e_p x_p = -\lambda_p e_p^2 \frac{\partial V}{\partial x_p} \quad \text{.....................(16).}$$

In the absence of an electric field the medium is unpolarized and the potential due to the distribution of charges forming the atomic systems is zero. Thus the potential V_p due to the polarized medium is that which arises from the totality of the doublets which correspond to (16).

the external field may change, the force acting on any assigned electron will always be changed in the same proportion. This result may be represented by putting

$$-e_p\frac{\partial V}{\partial s_p} = L_p e_p (E + \alpha P) \quad \text{.................(23)},$$

where L_p is a constant characteristic of the pth class of electrons.

Comparing with p. 73 we see that

$$\nu\sum_{p=1}^{p=n}\lambda_p L_p e_p^2 (E + \alpha P) = P = (\kappa - 1)E,$$

so that

$$P = \frac{\nu\Sigma\lambda_p L_p e_p^2}{1 - \alpha\nu\Sigma\lambda_p L_p e_p^2}E \quad \text{.................(24)}$$

and

$$\frac{\kappa - 1}{\kappa + \frac{1-\alpha}{\alpha}} = \frac{\Sigma\nu\lambda_p L_p e_p^2}{1 - \alpha\Sigma\nu\lambda_p L_p e_p^2}$$

$$\div\left(\frac{\Sigma}{1-\alpha\Sigma} + \frac{1}{\alpha}\left(\frac{1-\alpha\Sigma}{1-\alpha\Sigma}\right)\right)$$

$$= \alpha\nu\sum_{p=1}^{p=n}\lambda_p L_p e_p^2 \quad \text{....................(25)}.$$

λ_p, L_p and e_p may vary for different electrons in the same molecule but they will have the same value for corresponding electrons in different molecules of the same substance. The expression on the right hand side may therefore be represented by a summation over each molecule multiplied by the number of molecules of the substance in unit volume. We therefore find that

$$\frac{\kappa - 1}{\kappa + \frac{1-\alpha}{\alpha}} = k\rho \quad \text{.......................(26)},$$

where k is a constant and ρ is the density of the substance.

We shall find that the coefficients λ_p have an important significance when we come to consider the phenomenon of optical dispersion. The investigation leading up to formula (25) will not apply to optical problems without modification, as the displacement of the electrons in such cases is not necessarily always in phase with the corresponding "force."

When we are dealing merely with the electrostatic behaviour of dielectrics we can afford to neglect the complications just alluded magnetization, which is the name usually given to the quantity in magnetism which corresponds with intensity of polarization in electrostatics. The magnetic potential at a point is the work done in bringing a unit pole from a point at an infinite distance up to the point in question.

Just as in electrostatics we have to deal with dielectric media, so in magnetism there are media which, without being themselves permanently magnetized, have the power of modifying the distribution of magnetic force in their neighbourhoods. There is therefore a vector called magnetic induction analogous to the electric induction D. We shall denote the magnetic induction by the symbol B and its components by B_x, B_y, B_z. The coefficient which corresponds to the dielectric coefficient κ we shall call the permeability and denote by μ.

The behaviour of different media towards magnetic force furnishes more variety than the corresponding electrostatic phenomena. For a few substances μ may have very large values, and usually these are the substances which are capable of being permanently magnetized. Since iron is the typical example of this class of substances they are often called ferromagnetic. For all other substances μ does not differ greatly from unity, on the electromagnetic system of measurement. It is found that, in addition to the substances for which μ is greater than unity, and whose behaviour is analogous to that of dielectrics, there is another large class for which μ is less than unity. The former are said to be paramagnetic and the latter diamagnetic. It is very probable that paramagnetism and diamagnetism arise from the operation of separate causes. We shall, however, postpone the consideration of the physical causes which underlie the varied magnetic behaviour of substances until a later chapter. At present we are concerned with the formal relationship between magnetism and electrostatics.

If $H = H_x, H_y, H_z$ is the magnetic force it follows from the foregoing considerations, combined with the results of preceding chapters, that the following propositions are true for magnetism:

(1) The force between two poles of strengths m, m' embedded in a medium of permeability μ at a distance r apart is

$$F = \frac{mm'}{4\pi\mu r^2} \quad \text{.........................(1).}$$

surface, F_x, F_y, F_z are *actually* the components of the induction B_x, B_y, B_z. It follows that over any closed surface in space

$$\iint (lB_x + mB_y + nB_z)\, dS = 0 \quad\text{..............(14).}$$

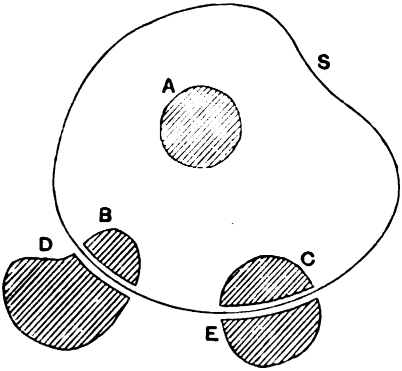


Fig. 16.

It follows that $$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0 \quad\text{....................(15)}$$

and the magnetic induction B is solenoidal *everywhere.* The corresponding result only holds for the electric induction in regions which do not contain any electric charges. Where there are "true" charges of volume density ρ

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} = \rho.$$

Force on a Magnetic Shell.

We have seen that the potential energy of a uniform magnetic shell of strength ϕ is $-N\phi$, where N is the number of lines of magnetic force threading it in the positive direction. This result enables us to calculate the force on the shell in the field, and we shall see that it may be represented as so much per unit length of the boundary of the shell. Let the shell be displaced so that N becomes $N + \delta N$. The diminution of the potential energy due to the displacement will be $\phi \delta N$, and this will be equal to the work done on the shell by the magnetic forces. The value of δN is easily calculated. Consider the prismatic figure, bounded by the original and displaced positions of the shell, which is traced out by the edge of the shell during the displacement. We shall suppose the displacement to be small and rectilinear; let us denote it by δx. Since there are no magnetic charges inside this figure the total flux of force over its boundary is zero. Hence

$$N + \delta N - N + \iint H \cos \widehat{nH} \sin \widehat{xs}\, ds\, \delta x = 0,$$

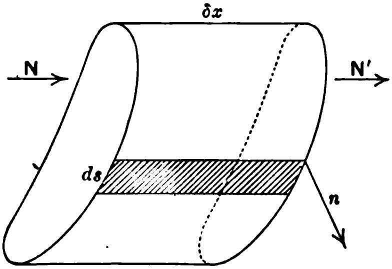


Fig. 17.

where H is the magnetic force at any point s of the boundary, $\widehat{xs}$ is the angle between the direction of the displacement δx and the element ds of the boundary, and $\widehat{nH}$ is the angle between the direction of H and the normal to the element of area $ds\,\delta x \sin \widehat{xs}$.

Thus $$\delta N = - \delta x \int H \cos \widehat{nH} \sin \widehat{xs}\, ds.$$

But if X is the force acting on the whole shell in the direction δx

$$X \delta x = \phi\, \delta N,$$

and $$X = - \phi \int H \cos \widehat{nH} \sin \widehat{xs}\, ds.$$

Thus the force X is equivalent to a set of forces of amount

$$- \phi H \cos \widehat{nH} \sin \widehat{xs}$$

per unit length of the edge of the shell. Since x is arbitrary this gives the force in any direction. To find the direction of the resultant force we notice that when x and s are in the same direction $\sin \widehat{xs} = 0$, so that the resultant force is at right angles to ds. It is also normal to H, since it vanishes when $\widehat{nH} = \pi/2$. It is therefore along the common normal to H and ds. To find the magnitude of this resultant let us suppose that the displacement δx is in the direction of this common normal. Then $\widehat{xs} = \pi/2$

lies. The potential must therefore have the same value after it is reversed as before. Its value must therefore be zero.

Now if *ABCD* is replaced by a coplanar circuit of any shape, by drawing a large number of radii from *O* it can be replaced by equivalent circuits bounded by radii and its original contour. The contour can be replaced by a series of concentric arcs since the original circuit will give rise to the same field if it is caused to zig-zag about its original path. Each of these constituent circuits gives rise to zero potential at the point *O*, so that the whole circuit gives rise to zero potential at *O*.

Now consider any circuit *ABCD* (Fig. 19) lying in space. Describe the cone on which the circuit lies and whose apex is at *O*. Also describe a sphere of unit radius about *O*. Let this cut the cone in the curve *EFGH*. Let i be the current round *ABCD*. From *O* draw a series of lines *OEA*, *OFB*, etc. intersecting the curves in *E*, *F*, *A*, *B*, etc. Imagine equal and opposite currents i to flow along each of these lines. Suppose also that a current i flows around the curve *GFEH* in the direction from *F* to *E*. This is the only new current that has been added; since the equal and opposite currents along *FB*, etc. cancel each other. The system now contemplated resolves itself into a series of currents i round circuits such as *ABFE*. Since *O* lies in the plane of each of these circuits the potential at *O* due to the system of currents is zero. The potential at *O* due to i round $F \rightarrow EHG$ is thus equal and opposite to that round $A \rightarrow BCD$. The potential due to i round

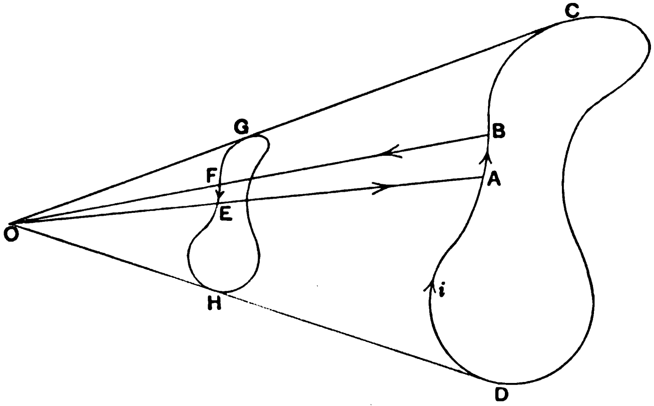


Fig. 19.

$A \rightarrow BCD$ is therefore the same as that due to the same current round $E \rightarrow FGH$. Now let us fill up the whole of the surface, cut off from the sphere of unit radius by the curve $EFGH$, with equal contiguous and superposable small areas. Imagine a current equal to i to flow round the boundary of each area. This will leave the original current i flowing round the boundary, and two equal and opposite currents along every line inside it. The potential due to the original circuit will therefore be equal to the potential due to the sum of the small circuits. Since all the small circuits are equal and have the same geometrical relation to O, the potential due to each one of them must be the same. It must be proportional to the current i since the force due to any system of currents changes in the same ratio as the currents when all the currents are changed in equal proportion. The argument is independent of the shape of the small circuits, and the only other condition to be satisfied is that the potential due to the whole circuit is the sum of those due to its parts. It follows that the geometrical factor to which the potential of each constituent circuit is proportional is $\delta\omega$, the solid angle it subtends at O. We therefore conclude that the potential due to each elementary circuit is $Ai\delta\omega$, where A is a constant which depends neither on the magnitude of the current nor on the geometry of the circuit.

It follows that when the current subtends a finite solid angle ω the potential to which it gives rise is $Ai\omega$. This is universally true since we have proved that all circuits carrying equal currents and lying on the same cone whose apex is at O, give rise to equal potentials at O. The magnitude of the constant A depends on the unit in which we measure the current. The usual electromagnetic unit of current is defined by making $A = 1$. In the units used in this book $A = (4\pi c)^{-1}$, where c is the velocity of light. (See later, p. 112.)

In our study of the properties of polarized shells we saw that the potential of a shell of strength ϕ was equal to $\phi\omega/4\pi$, where ω was the solid angle subtended by the shell at the point where the potential was measured. An electric current is thus equivalent in its magnetic action to a magnetic shell whose strength is equal to $4\pi A$ times the intensity of the current.

There is an important difference between the field due to an electric current and that arising from the equivalent magnetic shell. In the latter case the work done in taking a unit pole round a closed curve which passes once through the shell is zero, since the work done against the forces outside the shell is just equal to the work done by the forces inside the shell. In the case of the electric current there is nothing which fixes the position of the equivalent shell in space except the current which forms its boundary. It follows that there cannot be any region in which the force has the peculiar distribution characteristic of the interior of an actual magnetic shell. The force due to a current will therefore be continuous everywhere, and the work done in going round any closed path which embraces the current once will be equal to $4\pi Ai$. If the path followed by the unit pole circulates round the current m times in the positive sense and n times in the negative sense, the work done on it will be $4\pi A(m-n)i$. Thus the magnetic potential due to a current is determined not solely by the relative position of the current and the point, but also by the number of times the path of the point has previously encircled the current circuit. It is what is known as a multiple-valued function of space. The complete expression for the magnetic potential at any point due to a current may be written

$$Ai\,[4\pi(m-n)+\omega].$$

The force due to a current does not, of course, depend on m or n but only on ω, so that it will be single-valued and will depend only on the intensity and geometrical distribution of the current and the position of the circuit relative to the point.

The foregoing result that the work done in taking a unit magnetic pole once round any closed path embracing a current is proportional to the current embraced is not confined to linear currents, but is true if the currents occupy a finite volume. This is obvious if we divide up the whole current i into linear constituents δi bounded by tubes of flow. Then the work done in taking a unit pole round a path enclosing one of the constituents will be $4\pi A\delta i$ and if the path encloses the whole current i it will be

$$4\pi A\,\Sigma\,\delta i = 4\pi Ai \quad\text{.......................}(1).$$

The consideration of non-linear currents enables us to express this result in rather different analytical form. Consider any surface S in space, traversed by currents and bounded by a closed distribution of electric currents. They are one form of the first of Maxwell's two famous electromagnetic equations. To distinguish them from the second equation we shall sometimes refer to them as the statement of the First Law of Electrodynamics.

Like many of the formulae with which we have already become familiar, these three equations represent relations between components and derivatives of vectors. We shall frequently find it convenient to represent them by the abbreviated notation

$$\operatorname{rot} H = 4\pi A j,$$

where by rot H we mean the vector whose components are the quantities on the left hand side of equations (4), taken in succession. This equation is a vector equation, that is to say, it holds independently for each of the components of rot H and j respectively.

The Electric Current.

The electron theory regards the electric current which flows along a wire as a convection current. It supposes that in a conductor there are a number of charged particles which do not execute small displacements about a position of equilibrium, as in the case of those electrons which we considered when we were discussing the behaviour of dielectric media, but which are able to move freely from one part of the conductor to another. In a metallic conductor these particles are believed to be electrons and are called "free electrons" to distinguish them from the bound electrons, which only undergo small displacements from their position of equilibrium when an electric field is made to act on them. The free electrons in a metal are believed to be in much the same condition as the molecules of a gas. When we come, later on, to consider the evidence for this belief, we shall see that it is very strong and that the resemblance is a very close one. In the case of liquid electrolytes the charged particles are of atomic or molecular dimensions, and in many cases of electric conduction through gases this is the case also.

In all these cases the charged particles, of whatever nature, are believed to be moving about irregularly in all directions, even when they are not subjected to the action of an electric field. This motion does not cause any transportation of electricity since of the condenser B. This current is not, however, continuous all the way round the circuit, in the same way that it would be if the condenser were replaced by a wire, since there is no actual transportation of electricity across the dielectric between the plates of the condenser. Maxwell put forward the hypothesis that the magnetic effect of such a circuit was identical with that arising when the condenser is replaced by a suitable conductor, and there is the same current flowing along the wires in the circuit. This comes about owing to the existence, when the field is changing, of what Maxwell called a displacement current in the dielectric. The displacement current density at any point is equal to $\frac{dD}{dt}$, where D is the electric induction.

This value for the displacement current makes the current round the circuit continuous. For if σ is the charge per unit area of the plates of the condenser at any point, the current i in the wires is equal to $\iint \frac{d\sigma}{dt} dS$, the integral being taken over a plate of the condenser. But, if D is the value of the induction, the displacement current close to the plates is $\iint \frac{dD}{dt} dS$ taken over the same surface. We have seen that $D = \sigma$, so that the displacement current is equal to i and the current is continuous all round the circuit.

This result may be proved to be quite general as follows:

Consider any closed surface in space. By Gauss's Theorem the charge e inside this surface is equal to

$$\iint (lD_x + mD_y + nD_z)\, dS.$$

The current flowing into the surface is

$$\frac{de}{dt} = \iint \left(l\frac{\partial D_x}{\partial t} + m\frac{\partial D_y}{\partial t} + n\frac{\partial D_z}{\partial t} \right) dS.$$

But if j_x, j_y, j_z are the components of the true current density

$$\frac{de}{dt} = -\iint (lj_x + mj_y + nj_z)\, dS.$$

Hence

$$\iint \left[l\left(j_x + \frac{\partial D_x}{\partial t}\right) + m\left(j_y + \frac{\partial D_y}{\partial t}\right) + n\left(j_z + \frac{\partial D_z}{\partial t}\right) \right] dS = 0 \quad \ldots(6).$$

So that the vector $i = j + \frac{\partial D}{\partial t}$ is solenoidal everywhere. Like the tubes of magnetic induction, the tubes of flow of the total current are always closed regions of space. They have no free ends, unless at infinity.

It would be very difficult to devise a direct experimental proof of the magnetic effect of displacement currents. The best proof of it is an inductive one. The hypothesis of displacement currents forms the basis of the electromagnetic theory of light, and the extraordinary way in which the requirements of this theory have been fulfilled by experiment shows that it is built upon a solid foundation.

It is often desirable to consider the displacement current density $\frac{dD}{dt}$ as made up of two parts, (1) the aethereal displacement current $\frac{dE}{dt}$ and (2) the polarization current $\frac{dP}{dt}$. Since $D = E + P$, $\frac{dD}{dt}$ is always equal to the sum of the aethereal current and the polarization current. On the electron theory, as is obvious from the discussion in Chapter V, the polarization current corresponds to an actual displacement of charged electrons, and is to that extent very similar to the true current.

Convection Currents.

The kinds of electric currents which flow in wires, electrolytic cells and so forth, and which are carried by extremely minute particles, are not the only ones whose magnetic effects can be detected. Rowland* showed in 1876 that an electrostatically charged disc when made to revolve at a sufficiently high speed affected a suspended magnet in the same way as a current flowing round the disc. Effects of this kind, that is to say, effects depending on the movement of electricity on a large scale, may easily be summarized by means of a very simple formula. Let ρ be the net volume density of the electrification at any point, and let its velocity be given in magnitude and direction by V. Then ρV is the current density of the electricity at the point in question. This expression may be made to include not only cases where the motion of a single charge or a small number of charges is contemplated, but also those like the flow of electricity along a wire in which the number of carriers is enormous. In such cases we have to take the average value of ρV, and it is quite clear that this is equal to the magnitude of the true electric current as specified on p. 97.

---

* *Ann. der Phys.* vol. CLVIII. p. 87 (1876).

When it is necessary or advisable to distinguish between the different kinds of electric current which may occur, we may write equations (4) in the form

$$\text{rot}\, H = 4\pi A \left(\rho V + \frac{dD}{dt}\right) \quad\text{..................(7).}$$

Where we are dealing with material systems we shall require the average, not the actual, values of these vectors just as in the theory of the behaviour of dielectric media.

Induction of Currents.

In an electrostatic field the work done in taking a unit charge round any closed path, to the point from which it started, vanishes. The electrostatic potential is a single-valued function of the space coordinates, and the electromotive force round any closed circuit is zero. This is no longer true in a region in which the magnetic induction is changing. Faraday* showed that when the magnetic field inside a closed conducting circuit was made to change, a current was caused to flow round the circuit. He also showed that the electromotive force round the circuit was equal to the rate of diminution of the flux of induction through the circuit in the positive direction, multiplied by a universal constant. This statement is very similar in form to that in which we expressed the First Law. If we put it into analytical form we shall see that the resemblance is very close indeed.

The electromotive force round any closed circuit is equal to the work done in taking a unit positive charge round the circuit. If s denotes length measured along the circuit this is

$$\int E \cos \widehat{Es}\, ds = \int (E_x dx + E_y dy + E_z dz),$$

---

* *Exp. Res.* § 116.

necessary to give them a name we shall refer to them as the expression of the Second Law of Electrodynamics. Comparing them with the First Law,

$$\operatorname{rot} H = 4\pi A\left(\rho V + \frac{\partial D}{\partial t}\right),$$

we see that, since B is proportional to H and D to E, there is a kind of reciprocal relation between them. Also they are not quite the same in form on account of the negative sign in front of $\frac{\partial B}{\partial t}$.

Not only do these equations enable us to deduce the whole of the phenomena of electromagnetism, but, as we shall see, they are the basis of the science of optics as well.

The Dynamical Theory of Electromagnetism.

The Second Law of Electromagnetism may be looked upon from two different standpoints according to the attitude we take towards electrical science. If we regard electrodynamics as more fundamental than dynamics proper, then we must regard the Second Law as a fundamental law of nature empirically given. We may however take the standpoint that the aether, which we postulate as a medium in which all electrical actions occur, will in the last analysis prove to be a mechanical system subject to the basal laws of dynamics. Provided we make this assumption, even though we know nothing of the nature of the mechanism, we can show that the Second Law is a consequence of the First Law. The view that electrical actions are ultimately dynamical is one whose development in the hands of Maxwell led to notable advances in the science, and it is the view towards which, at any rate until quite recently, most authorities have leaned. Nevertheless it is equally logical to accept the Second Law as an ultimate fact and then, later on, to consider what we can make of the laws of dynamics from the standpoint thus adopted.

We shall now proceed to consider some of the consequences which follow from the assumption that every electrodynamical system is a dynamical system subject to the operation of the first law and of the fundamental laws of electrostatics and magnetostatics. The energy of the field is equal to

$$\tfrac{1}{2}\kappa(E_x^2 + E_y^2 + E_z^2) + \tfrac{1}{2}\mu(H_x^2 + H_y^2 + H_z^2) \quad \ldots\ldots(9)$$

per unit volume. Since a change in the electric intensity is accompanied by a corresponding change in the magnetic force, it is natural to assume that one of these terms represents potential and the other kinetic energy. Moreover the presence of magnetic energy is always associated with charges in motion whereas electrostatic energy is present when all the charges are at rest. The obvious conclusion therefore is that magnetic energy should be identified with kinetic energy and electrostatic with potential energy. A more rigorous proof of the necessity for identifying magnetic with kinetic energy will be found in Jeans's *Electricity and Magnetism*, p. 483.

Assuming that the magnetic energy of the electrodynamic field represents the kinetic energy of a dynamical system, let us consider the behaviour of a system of n circuits 1, 2, 3, etc. carrying currents i_1, i_2, i_3, etc. In this case the magnetic energy can be written as a quadratic function of the n variables i_1, i_2, i_3, etc. For the number of tubes of magnetic induction which flow through any one, let us say the mth, circuit, is the sum of n terms each representing a contribution from one circuit. This follows since the magnetic force at any point due to a current depends only on the relative geometry of the point and the current, the nature of the intervening medium, and the magnitude of the current, to which it is proportional. Thus if N_m is the number of tubes of magnetic induction which traverse the mth circuit in the positive direction

$$N_m = L_{1m}i_1 + L_{2m}i_2 + L_{3m}i_3 + \dots + L_{mm}i_m + \dots + L_{nm}i_n \quad \dots(10).$$

The coefficients L_{1m}, etc. depend only on the geometry of the circuits and the nature of the material in which they lie. They are called coefficients of self-induction when the suffixes are like, and of mutual induction when they are unlike.

The magnetic energy of any system, including that of a system of currents, is equal to $\frac{1}{2}\iiint \mu H^2 d\tau$, taken throughout the volume of the system. Let us suppose that the space is mapped out by means of unit tubes of induction. These are closed tubes which completely fill the space and never intersect. If δS is the normal sectional area of a unit tube at any point, the element of volume $d\tau$ may be replaced by $\delta S\, ds$, where ds is an element of length of the tube and is normal to δS. But δS is the area over which the flux of induction is unity, so that

$$\mu H\, \delta S = 1.$$

Chap. II, p. 33, in obtaining the total energy of a system of electrostatic charges, except for the difference in the form of the expression for the potential energy of the system. This calculation introduces a factor $\frac{1}{2}$, so that the energy of the system of shells which produces the same magnetic field as the system of currents is

$$-2\pi A\,\Sigma i_m N_m = -2\pi A\,(L_{11}i_1^2 + 2L_{12}i_1 i_2 + \ldots + L_{nn}i_n^2).$$

The difference between this expression and (11) arises from the energy which lies within the volume of the shell itself*. This quantity, which is equal to $-2T$ in the case of the equivalent shells, is zero in the case of the currents.

The method of Lagrange's equations† enables us to find out a great deal about the behaviour of a mechanical system even when we have no means of discovering the precise nature of the mechanism. We shall therefore apply that method to the problem under consideration. As a preliminary we have to express the kinetic energy T and the potential energy W as functions of the generalized coordinates x_1, x_2, etc. and the corresponding velocities $\dot{x}_1$, $\dot{x}_2$, etc. We have seen that if we identify T with the magnetic energy in the field we get

$$T = +\,2\pi A\,(L_{11}i_1^2 + L_{12}i_1 i_2 + L_{13}i_1 i_3 + \ldots),$$

and this will be a quadratic function of the velocity coordinates $\dot{x}_1$, $\dot{x}_2$, etc. if we make $i_1 = \dot{x}_1$, $i_2 = \dot{x}_2$, $i_3 = \dot{x}_3$, etc. Thus the generalized displacement x_s becomes $\int i_s dt$ and is equal to the quantity of electricity which has flowed round the circuit after some fixed instant. In the present case $W = 0$ since the system does not possess electrostatic energy, the capacity being regarded as negligible. We also notice that the L's are functions only of the geometrical arrangement of the circuits and of the nature of the intervening medium. They do not involve the generalized displacements x.

If X_s is the component of generalized external force corresponding to the generalized displacement x_s, we have

$$\frac{\partial}{\partial t}\left(\frac{\partial T}{\partial \dot{x}_s}\right) - \frac{\partial T}{\partial x_s} = X_s \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots(12),$$

---

* Cf. Jeans, *Electricity and Magnetism*, p. 433.

† For an account of the part of generalized dynamics which is germane to the present discussion the reader may be referred to Jeans's *Electricity and Magnetism*, Chap. XVI.

First Law in the form rot $\frac{\partial\theta}{\partial t} = 4\pi A\frac{\partial D}{\partial t}$, and we shall extend this equation so as to make it true of any change of the dependent variables θ and D whether in respect to the time or not. Thus if θ varies by $\delta\theta$ we shall suppose the variation to be conditioned by a variation δD of D according to the formula rot $\delta\theta = 4\pi A\,\delta D$. We have, if $d\tau$ is an element of volume,

$$\begin{aligned}\delta\int d\tau\int_{t_1}^{t_2} T\,dt &= \int d\tau\int_{t_1}^{t_2}\mu(\dot\theta_x\delta\dot\theta_x + \dot\theta_y\delta\dot\theta_y + \dot\theta_z\delta\dot\theta_z)\,dt\\ &= \int\Big|\mu(\dot\theta_x\delta\theta_x + \dot\theta_y\delta\theta_y + \dot\theta_z\delta\theta_z)\Big|_{t_1}^{t_2}\times d\tau\\ &\qquad -\iint_{t_1}^{t_2}\{\mu(\ddot\theta_x\delta\theta_x + \ddot\theta_y\delta\theta_y + \ddot\theta_z\delta\theta_z)\}\,d\tau\,dt,\end{aligned}$$

and $$\delta\iint_{t_1}^{t_2} W\,d\tau\,dt = \int d\tau\int_{t_1}^{t_2}\kappa(E_x\delta E_x + E_y\delta E_y + E_z\delta E_z)\,dt$$

$$\begin{aligned}&= \frac{1}{4\pi A}\int d\tau\int_{t_1}^{t_2}\Big[E_x\Big(\frac{\partial}{\partial y}(\delta\theta_z) - \frac{\partial}{\partial z}(\delta\theta_y)\Big) + E_y\Big(\frac{\partial}{\partial z}(\delta\theta_x) - \frac{\partial}{\partial x}(\delta\theta_z)\Big)\\ &\qquad + E_z\Big(\frac{\partial}{\partial x}(\delta\theta_y) - \frac{\partial}{\partial y}(\delta\theta_x)\Big)\Big]\,dt\\ &= \frac{1}{4\pi A}\int_{t_1}^{t_2} dt\int dS\Big[\delta\theta_x(nE_y - mE_z) + \delta\theta_y(lE_z - nE_x)\\ &\qquad + \delta\theta_z(mE_x - lE_y)\Big]\\ &\quad - \frac{1}{4\pi A}\int_{t_1}^{t_2} dt\int d\tau\Big[\delta\theta_x\Big(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y}\Big) + \delta\theta_y\Big(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z}\Big)\\ &\qquad + \delta\theta_z\Big(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}\Big)\Big].\end{aligned}$$

Thus

$$\begin{aligned}&\delta\int d\tau\int_{t_1}^{t_2}(T - W)\,dt = 0 = \int d\tau\,\Big|\mu(\dot\theta_x\delta\theta_x + \dot\theta_y\delta\theta_y + \dot\theta_z\delta\theta_z)\Big|_{t_1}^{t_2}\\ &\quad - \frac{1}{4\pi A}\int d\tau\int_{t_1}^{t_2}\Big\{\delta\theta_x\Big(4\pi A\,\mu\ddot\theta_x + \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\Big)\\ &\qquad + \delta\theta_y\Big(4\pi A\,\mu\ddot\theta_y + \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\Big)\\ &\qquad + \delta\theta_z\Big(4\pi A\,\mu\ddot\theta_z + \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\Big)\Big\}\,dt\\ &\quad - \frac{1}{4\pi A}\int dS\int_{t_1}^{t_2}\Big[\delta\theta_x(nE_y - mE_z) + \delta\theta_y(lE_z - nE_x)\\ &\qquad + \delta\theta_z(mE_x - lE_y)\Big]\,dt.\end{aligned}$$

The present analysis is restricted to the case where the media are at rest and where there are no free charges. Moreover the actual and the varied system are to be identical at $t = t_1$ and $t = t_2$. This makes the first volume integral vanish, since at both time limits $\delta\theta_x = \delta\theta_y = \delta\theta_z = 0$. The surface integral will vanish if we suppose the surface to be a closed surface at an infinite distance from the region in which the actions are going on. For although dS is proportional to r^2 both E and θ are of the order $\frac{1}{r^2}$ at most. It follows that the second volume integral must vanish. And since the $\delta\theta$'s are perfectly arbitrary both for each element of volume and for each element of time (except for the limits of time) the coefficients of each of them must always be zero. We therefore find

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -4\pi A\mu\ddot{\theta}_x = -4\pi A\frac{\partial B_x}{\partial t},$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -4\pi A\mu\ddot{\theta}_y = -4\pi A\frac{\partial B_y}{\partial t},$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -4\pi A\mu\ddot{\theta}_z = -4\pi A\frac{\partial B_z}{\partial t},$$

or

$$\operatorname{rot} E = -4\pi A\frac{\partial B}{\partial t}.$$

Thus the Second Law follows if we assume the truth of the First Law, and make the further assumption that the electrodynamic field is a mechanical system.

The foregoing deduction is practically identical with one given by Larmor (*Aether and Matter*, Chapter VI), who has shown that the laws of electrodynamics can be built up by giving the aether a mechanical constitution. For the further development of this theory, including the natural extension to the case where the presence in the field of electric charges is contemplated, the reader may be referred to Larmor's *Aether and Matter*, Chaps. VI, VII, and X.

Electrical Units.

We have seen that the work done in taking a unit magnetic pole round a path situated in free space which encircles a current of strength i once, is given by

$$\int H\,ds = 4\pi Ai.$$

The value of the universal constant A will clearly depend on the magnitude of the units in which H and i are measured. It also depends on the units of length and time.

We have already defined our units of magnetic pole strength and electric charge. From either of these units, although not from both, the units of all other electric and magnetic quantities may be derived unequivocally provided we keep to the same units of mass, length and time. Thus, to determine our unit of magnetic intensity H we only need to remember that mH is the force in dynes on a pole of strength m, so that the unit of H is determined by the unit of m and of force. It is evident that the dimensions of the product mH are independent of the dimensions which, since they are unknown, may arbitrarily be assigned to m; the product must always have the dimensions of a mechanical force. The dimensions of many other combinations of electric and magnetic quantities are predetermined in the same way; for instance, Ee has the dimensions of force and μH^2 and κE^2 have the dimensions of energy per unit volume, and so on. The unit of current i clearly only involves the units of electric charge and time, and is therefore determined in our case since we have already fixed the units of electric charge and time.

When H and i are measured in this way, the value of the constant factor is $4\pi A = 1/c$, where $c = 3 \times 10^{10}$ cms. per sec. The quantity c, which has the dimensions of a velocity, and thus has a numerical value which depends only on the units of length and time, is one of the most important physical quantities. As we shall see, it is equal, among other things, to the velocity of light and other forms of electromagnetic radiation in empty space.

When the stipulation above as to the character of the derived units is understood we can write equation (7) in the form

$$\text{rot}\, H = \frac{1}{c}\left(\frac{\partial D}{\partial t} + \rho V\right).$$

It remains to consider the value of the constant A_1 which enters into the expression for the second law. This may be discovered by making use of the principle of the conservation of energy, and for this purpose the simplest possible example of the induction of currents will suffice as well as another. Consider a single circuit carrying a current i, the self-induction of the circuit being L. Then the magnetic or kinetic energy of this circuit is $2\pi A L i^2 = \frac{L}{2c} i^2$. If the current i changes, the amount of this energy will alter and it is a known physical fact that there will be in consequence a self-induced electromotive force acting round the circuit. In general the circuit may lose energy by radiation or in other ways, but under suitable conditions these losses are negligible. In such cases the rate of loss of magnetic energy will be equal to the work done against the resistance of the circuit less the work supplied by the battery of electromotive force E. Thus

$$\frac{1}{2c}\frac{\partial}{\partial t}(Li^2) = Ei - Ri^2,$$

or

$$\frac{L}{c}\frac{\partial i}{\partial t} = E - Ri.$$

The electromotive force due to induction is therefore

$$E' = -\frac{L}{c}\frac{\partial i}{\partial t} = -\frac{1}{c}\frac{\partial N}{\partial t},$$

where N is the number of tubes of induction, measured in our system of units, which thread the circuit in the positive direction. It follows by comparison with the equations on p. 101 that $A_1 = \frac{1}{c} = 4\pi A$, when the electrostatic quantities are measured in the modified electrostatic units and the magnetic quantities in the modified electromagnetic units.

The following example is also instructive, since it brings out quite clearly that it must be the magnetic induction and not the magnetic intensity whose rate of change determines the magnitude of the induced currents. Consider the behaviour of a bar of magnetizable material encircled by a solenoid having n turns of wire per unit length. The bar is of uniform cross section a, and of indefinite length l. A current i flows in the solenoid and maintains a magnetic field of intensity H in the bar. The relation between H and i is

$$H = 4\pi A n i = \frac{n}{c} i \quad \dots\dots\dots\dots\dots\dots\dots (16).$$

The work done in establishing the magnetic field in the bar consists of two parts, (1) $la\int H dI$, which, if I is the intensity of magnetization, represents the actual work in magnetizing the bar, and (2) $la\int H dH$, which represents the energy of the magnetic field in the aether of the space occupied by the bar. In a small interval of time dt the increment in the magnetic energy in the bar is thus

$$laH\left(\frac{dI}{dt}+\frac{dH}{dt}\right)dt.$$

This must be equal to the work done against the back electromotive force of induction. This work is

$$A_1 nla\frac{dB}{dt}\,.\,idt = laH\,.\,A_1 c\,.\,\frac{dB}{dt}\,dt.$$

So that
$$\frac{d}{dt}(I+H)=A_1 c\frac{dB}{dt}.$$

Thus $B=H+I$ in accordance with Chap. VI and $A_1=\frac{1}{c}=4\pi A$.

It is well to understand clearly the difference between the system of units we are using and the two systems, the electrostatic system and the electromagnetic system respectively, which are most frequently used in books dealing with the theory of electricity. The unit of electric charge in the electrostatic system is $\sqrt{4\pi}$ times our unit of electric charge; but it is not this difference so much as the difference in the units in which the magnetic quantities are measured which it is desirable to emphasize at the moment. On the electrostatic system the unit of current is obtained from the unit of electric charge, and the magnetic quantities are then obtained by giving A the arbitrary value unity in the equation

$$\int H ds = 4\pi A i.$$

This fixes the unit of magnetic force and so determines the unit of magnetic charge. In our units the measure of i is $\sqrt{4\pi}$ times greater and A is $4\pi c$ times less than in the electrostatic system, so that our unit of magnetic force is $c\sqrt{4\pi}$ times greater than the electrostatic unit. Since mH has the same value on all systems which have the same unit of mechanical force our unit of pole strength is $c\sqrt{4\pi}$ times smaller than the electrostatic unit.

The electromagnetic system of units also makes $A=1$, but it sets out by defining the unit magnetic pole as that which repels an equal pole at unit distance with a force of one dyne. On this

CHAPTER VII

ELECTROMAGNETIC WAVES

The Equations of Propagation.

LET us for the moment confine our attention to the application of Maxwell's equations to regions in which there are neither electric nor magnetic charges. Strictly speaking, from the point of view of the electron theory, this should restrict us to the case of the free aether, since all matter is supposed to be made up of electrically charged particles. In material media, however, the mean density of the charge at any point, if we average over a volume containing an enormous number of particles, is in most cases zero; so that we shall examine the consequences, incidentally, of supposing that with material media the peculiarities of the individual electrons can be left out of consideration. We shall soon see that the results at which we arrive, while exact for the free aether, are the crudest kind of approximation when applied to material media.

The equations to which we have been led in the preceding chapter in the case in which $\rho = 0$ are

$$\operatorname{rot} H = \frac{1}{c}\frac{\partial D}{\partial t} = \frac{\kappa}{c}\frac{\partial E}{\partial t},$$

$$\operatorname{rot} E = -\frac{1}{c}\frac{\partial B}{\partial t} = -\frac{\mu}{c}\frac{\partial H}{\partial t}.$$

We have seen that the two foregoing equations are really an abbreviation for six equations between the six components of H and E. These equations can therefore be solved for each one of by Green's Theorem (Chap. II, p. 24), where the surface integral extends over the surface enclosing the volume. Let S be the surface of a sphere of radius r whose centre is at the point P, and let $d\omega$ be the solid angle subtended by dS at the point P, then

$$-\iint \frac{\partial u}{\partial n} dS = \iint \frac{\partial u}{\partial r} r^2 d\omega = r^2 \frac{\partial}{\partial r} \iint u_r d\omega,$$

and $$\iiint \frac{\partial^2 u}{\partial t^2} d\tau = \iiint \frac{\partial^2 u}{\partial t^2} r^2 d\omega\, dr = \frac{\partial^2}{\partial t^2} \iint d\omega \int_0^r r^2 u_r dr.$$

Differentiating both sides of (8) by the upper limit r,

$$r^2 \frac{\partial^2}{\partial t^2} \iint u_r d\omega = a^2 \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r} \iint u_r d\omega\right).$$

Now $\iint u_r d\omega$ is 4π times the mean value of u over the surface of the sphere whose centre is P and radius r. Let us denote this by $\overline{u_r}$, then

$$r^2 \frac{\partial^2}{\partial t^2} \overline{u_r} = a^2 \frac{\partial}{\partial r}\left(r^2 \frac{\partial \overline{u_r}}{\partial r}\right),$$

$\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r} \frac{v}{r}\right)$

or $$\frac{\partial^2}{\partial t^2}(r\overline{u_r}) = a^2 \frac{\partial^2}{\partial r^2}(r\overline{u_r}).$$

$= a^2$ [illegible]

Now let $r\overline{u_r} = v$ and introduce new independent variables $p = at + r$ and $q = at - r$, then

$$\frac{\partial v}{\partial t} = \frac{\partial v}{\partial p}\frac{\partial p}{\partial t} + \frac{\partial v}{\partial q}\frac{\partial q}{\partial t} = a\left(\frac{\partial v}{\partial p} + \frac{\partial v}{\partial q}\right),$$

$\frac{\partial}{\partial t} = a\left(\frac{\partial}{\partial p} + \frac{\partial}{\partial q}\right)$

$$\frac{\partial^2 v}{\partial t^2} = a^2\left(\frac{\partial^2 v}{\partial p^2} + 2\frac{\partial^2 v}{\partial p \partial q} + \frac{\partial^2 v}{\partial q^2}\right),$$

$$\frac{\partial v}{\partial r} = \frac{\partial v}{\partial p}\frac{\partial p}{\partial r} + \frac{\partial v}{\partial q}\frac{\partial q}{\partial r} = \frac{\partial v}{\partial p} - \frac{\partial v}{\partial q},$$

$$\frac{\partial^2 v}{\partial r^2} = \frac{\partial^2 v}{\partial p^2} - 2\frac{\partial^2 v}{\partial p \partial q} + \frac{\partial^2 v}{\partial q^2},$$

so that $$\frac{\partial^2 v}{\partial p \partial q} = 0.$$

The general solution of this equation is clearly

$$\begin{aligned} v &= f_1(p) + f_2(q) \\ &= f_1(at + r) + f_2(at - r) = r\overline{u_r}, \end{aligned}$$

where f_1 and f_2 are arbitrary functions.

Since $\overline{u_r}$ is never infinite, $v = 0$ when $r = 0$, so that

$$f_2(at) = -f_1(at) \text{ for all values of } t.$$

The value of u_P will then be determined, provided we know the average value of $\frac{\partial u}{\partial t}$ over the surface of this sphere at $t=0$, as well as the rate at which the average value of u over this sphere at time $t=0$ changes as the radius is increased.

The meaning of this solution is perhaps seen more clearly if we consider a case in which the value of u at $t=0$ is zero everywhere except in a certain limited region of space τ. Consider the value of u at some external point P. Let the shortest line from P to τ be b and the longest line c. Then until $t=b/a$, the value of u_P will be zero since the sphere of radius at will not intersect the region for which u was not equal to zero at $t=0$. For similar reasons when t is greater than c/a the value of u_P will again be zero. Thus u represents a disturbance which is propagated in all directions with the velocity a. The foregoing solution, which is due to Poisson, shows that electromagnetic effects are propagated with finite velocity, like waves. In fact the equation

$$\frac{\partial^2 u}{\partial t^2}=a^2\nabla^2 u$$

is the general equation of wave motion and contains the mathematics underlying all the different kinds of wave motion contemplated in physics.

Before leaving Poisson's solution it may be well to consider a very simple concrete example to which it may be applied. Suppose that by means of two current sheets AB and CD perpendicular to the plane of the paper, we produce a uniform magnetic field H in the direction of the arrow. Suppose that at a certain time $t=0$ the currents are stopped. What will be the value of the magnetic intensity at the point P afterwards? Clearly there will be no field at P until $t=\frac{PR}{a}=\frac{b}{a}$, where a is the velocity of propagation; that is to say, until the sphere of radius at cuts the plane CD. To find the value after the sphere has

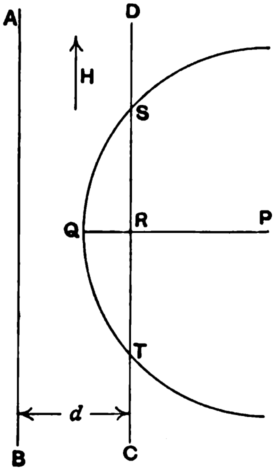


Fig. 22.

cut the plane CD but not the plane AB, let the points of intersection with CD be S and T. Then at $t=0$, $\frac{\partial H}{\partial t}=0$ everywhere except in a circular strip of infinitesimal width of which ST is the diameter. Even in this infinitesimal strip $\frac{\partial H}{\partial t}$ is finite, so that the integral in (10) which involves F' is zero when $t=0$. Also $H=\text{constant}=F$ at every point of the spherical cap of which SQT is a section, and $\frac{1}{4\pi}\iint F\,d\omega=\frac{H\times QR}{2PQ}=H\,\frac{at-PR}{2at}$. Thus from equation (10) $H_P=\frac{H}{2}$, and is independent of the position of Q so long as it lies between the planes. When the sphere of radius at intersects both planes $\frac{1}{4\pi}\iint F\,d\omega=\frac{H}{2}\frac{d}{at}$, where d is the distance between the planes, and

$$H_P=\frac{\partial}{\partial\,(at)}\left(\frac{at}{4\pi}\iint F\,d\omega\right)=0.$$

It is clear that the slab of thickness d in which the magnetic intensity is H splits into two slabs of equal thickness in which the magnetic intensity is $H/2$, and these are propagated in opposite directions normal to the faces of the slab with the velocity a.

The moving slabs of magnetic force form only part of the solution of the problem. This is clear because the energy of the magnetic field in the two moving slabs taken together is only half of the energy of the original magnetic field. It is necessary to consider the electric intensity E as well.

At $t=0$, $E=0$ everywhere, but $\frac{\partial E}{\partial t}$ is not zero. Consider any rectangle δx, δy in the plane of the paper such that δx crosses CD to which δy is parallel. Then, by the first law of electrodynamics, confining ourselves to the case of propagation in free space for which $\kappa=1$, $\mu=1$ and $a=c$,

$$\frac{1}{a}\iint\frac{\partial E}{\partial t}\,\delta x\,\delta y=\int H\,ds,$$

or

$$\frac{1}{a}\int\frac{\partial E}{\partial t}\,\delta x=H,$$

since everything is independent of y and the work $H\delta y$ vanishes along the right-hand side of the rectangle where $H = 0$. But if the angle $SPQ = \theta$,

$$\frac{\partial E}{\partial t}\,d\omega = \frac{\partial E}{\partial t}\frac{2\pi at\sin\theta\, atd\theta}{a^2t^2} = -2\pi\frac{\partial E}{\partial t}\,d(\cos\theta) = \frac{2\pi}{at}\frac{\partial E}{\partial t}\,dx.$$

Thus $\iint \frac{\partial E}{\partial t}\,d\omega = \frac{2\pi}{at}\int\frac{\partial E}{\partial t}\,\delta x$ and, from (10),

$$E_P = 0 + \frac{t}{4\pi}\iint\frac{\partial E}{\partial t}\,d\omega = \frac{1}{2a}\int\frac{\partial E}{\partial t}\,dx = \frac{H}{2}.$$

When the sphere of radius PQ cuts both planes we see that the value of $\int\frac{\partial E}{\partial t}\,\delta x$ contributed by the second plane is equal and opposite to that from the first, so that E_P again becomes zero. Thus the slab of moving magnetic force H is accompanied by an equal electric force at right angles to and of equal magnitude with H. The electric force in the slab which moves to the right is in the opposite direction to that in the slab which moves to the left, whereas the magnetic force is in the same direction in both cases.

Velocity of Propagation.

We have seen that the components of the electric and magnetic intensities satisfy the equation $\frac{\partial^2 u}{\partial t^2} = a^2\nabla^2 u$, so that their changes in time and space are such as would arise if they were propagated from every point with velocity $V = c/\sqrt{\mu\kappa}$. For a vacuous space $\mu = 1$ and $\kappa = 1$, so that it follows that electromagnetic disturbances are propagated in vacuo with a velocity which is equal to c, the ratio of the electromagnetic unit of charge to the electrostatic unit. At the time that this conclusion was first reached (by Maxwell) it had not been shown that electromagnetic disturbances were propagated with finite velocity, so that there was no experimental material available, by means of which the conclusion could be quantitatively tested. Maxwell however put forward the view, which Faraday's instinct had previously led him to express although his mathematical limitations had probably prevented him from being able to deduce adequate experimental evidence in favour of it, that light itself was really an electromagnetic phenomenon. In that event, Maxwell predicted, its velocity of propagation should be equal to the ratio c of the two kinds of units. Maxwell himself carried out a determination of the value of c and found it to be equal to the velocity of light through space, within the limits of his experimental error. At a much later date (1888) Maxwell's Theory was put in a much stronger position by the researches of Hertz* who showed that vibrating electrical systems such as could be set up in the laboratory emitted electromagnetic waves. Hertz investigated the principal properties of these waves and showed that they were analogous to waves of light, from which they differed principally in the possession of much longer wave-lengths.

More recent determinations have shown a continually increasing agreement between the values of (1) the ratio of the two units of electric charge, (2) the velocity of electric waves and (3) the velocity of light in space. The agreement of the results of different observers is well exhibited by the following numbers which are taken from Jeans's *Electricity and Magnetism*, p. 506.

For the value of c, the ratio of the two units, the following results have been collected by H. Abraham† as likely to be most accurate:

Himstedt	$3·0057 \times 10^{10}$	Abraham	$2·9913 \times 10^{10}$
Rosa	$3·0000 \times 10^{10}$	Pellat	$3·0092 \times 10^{10}$
J. J. Thomson	$2·9960 \times 10^{10}$	Hurmuzescu	$3·0010 \times 10^{10}$
Perot and Fabry	$2·9973 \times 10^{10}$		

The mean of these quantities is

$$c = 3·0001 \times 10^{10} \text{ cms./sec.}$$

For the velocity of propagation of electromagnetic waves in air the following values are collected by Blondlot and Gutton‡:

Blondlot	$3·022 \times 10^{10}$,	$2·964 \times 10^{10}$,	$2·980 \times 10^{10}$
Trowbridge and Duane			$3·003 \times 10^{10}$
MacLean			$2·9911 \times 10^{10}$
Saunders		$2·982 \times 10^{10}$,	$2·997 \times 10^{10}$

The mean of these quantities is

$$2·991 \times 10^{10} \text{ cms./sec.}$$

---

* *Ann. der Phys.* vol. xxxiv. p. 551 (1888).

† *Rapports du Congrès de Physique*, Paris, 1900, vol. II. p. 267.

‡ *Rapports du Congrès de Physique*, Paris, 1900, vol. II. p. 283.

electrons is much the same as it would be under a steady field of the same magnitude as the instantaneous value of that due to the light wave. Under these circumstances the material can be treated as a continuous medium of definite dielectric coefficient κ, and for waves of this character the velocity of propagation in different media should be inversely as the square roots of the dielectric coefficients. This conclusion is substantiated by the results of experiments. Thus A. D. Cole[*] found the refractive index of water to be 8·9 whereas its dielectric coefficient $\kappa = 80$. With other substances the agreement appears to be satisfactory within the rather considerable limits of error of the determinations of the dielectric coefficients[†].

We shall defer to the next chapter the consideration of the causes which make the behaviour of bodies towards light different from that predicted by the simple form of the electromagnetic theory which we have been discussing. There are, however, a number of phenomena exhibited by electromagnetic waves in their relation to matter which are partly true for light waves and strictly true for very long waves. The rest of this chapter will be occupied with an account of some of these.

Properties of a Plane-Polarized Electromagnetic Wave.

A solution of the equation $\frac{\partial^2 u}{\partial t^2} = a^2 \nabla^2 u$ is

$$u = u_0 e^{i \frac{2\pi}{\lambda}(at - lx - my - nz)} \quad \text{.................}(11)$$

provided $l^2 + m^2 + n^2 = 1$. The expression on the right-hand side is a complex quantity, being equal to

$$u_0 \left[\cos \frac{2\pi}{\lambda}(at - lx - my - nz) + i \sin \frac{2\pi}{\lambda}(at - lx - my - nz)\right].$$

The real part of u therefore represents a disturbance of wave-length λ and amplitude u_0 which is propagated along the straight line $x/l = y/m = z/n$ with constant amplitude u_0 and constant velocity a. It is thus the appropriate specification of a monochromatic train of plane waves of wave-length λ. If we take the direction of propagation to be along the axis of z we shall have $n = 1$ and u will be equal to $u_0 e^{\frac{2\pi}{\lambda}(at-z)}$ and its real part to

---

[*] *Wied. Ann.* vol. LVII. p. 290 (1896).

[†] Fleming, *Principles of Electric Wave Telegraphy*, p. 320.

$$u_0 \cos \frac{2\pi}{\lambda}(at-z) \quad \text{....................(12).}$$

Let the real part of u represent E_x, the x component of the electric intensity in the wave front, and consider the train of waves for which $E_y = E_z = 0$. The electric intensity in this train will be completely specified by the equations

$$E_x = \text{real part of } E_0 e^{i\frac{2\pi}{\lambda}(at-z)}$$

or

$$E_x = E_0 \cos \frac{2\pi}{\lambda}(at-z) \quad \text{..................(13).}$$

It is clear that any equation between functions of complex variables such as, for example,

$$F_1(u, iv) = F_2(x, iy)$$

involves the separate truth of the two equations

$$\text{Real part of } F_1 = \text{Real part of } F_2$$

and

$$\text{Imaginary part of } F_1 = \text{Imaginary part of } F_2,$$

otherwise $i = \sqrt{-1}$ would be equal to a real quantity, which is absurd. This principle effects considerable simplification in the working out of problems arising in connection with the propagation of waves, as it enables us to work with the complex solution and then pick out the real parts at the end of our calculations. The advantage of this lies in the fact that the complex equations are usually simpler than their real equivalents.

Suppose that we are dealing with the train of plane waves propagated along the axis of z. Each of the vectors E_x, E_y, E_z, H_x, H_y, H_z which serve to specify the state of the medium at any point at any instant must be of the form $u_0 e^{i\frac{2\pi}{\lambda}(at-z)}$. The values however are not independent but have to satisfy the six equations on p. 116, viz.

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = \frac{\kappa}{c}\frac{\partial E_x}{\partial t},$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = \frac{\kappa}{c}\frac{\partial E_y}{\partial t},$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = \frac{\kappa}{c}\frac{\partial E_z}{\partial t},$$

to two planes at right angles to each other passing through the direction of propagation. We shall see that the plane which is usually called the plane of polarization in optics is that which contains the direction of the magnetic vector and is perpendicular to the direction of the electric vector. (See p. 132.)

It is evident that as the wave moves past any fixed point O, the electric and magnetic intensities always remain proportional to one another and both vanish and also both have their maximum values simultaneously.

In the case of waves of sound the kinetic energy is measured by the squares of the time rate of change of the displacement of the particles constituting the medium, and the potential energy by the squares of the strains which depend upon the rate of change of the same displacements in space. Both the velocities and the strains travel together or are in phase just as the magnetic energy and the electric energy in electromagnetic waves are in phase. This may be regarded as another reason for identifying magnetic and electric energy with kinetic and potential energy respectively.

The following elementary method of deducing the velocity of an electromagnetic wave is instructive. Let the wave be propagated along the line Oz and let E_x be the x component of the electric intensity in the wave front. Consider the state of things when the wave front lies in the plane perpendicular to Oz which passes through the point O. Then the lines OQ and OU (Fig. 23) are in the wave front. Describe the rectangles $TOPRUS$ and $TOPVQW$, TP being small compared with PR and PV. TP, SR and WV are perpendicular to the wave front, PR and ST are parallel to OU, the axis of x, and PV and TW are parallel to OQ, the axis of y. Consider the work done in taking a unit magnetic pole round the rectangle $PTWV$. The work along the part $QVPO$ of the path vanishes; for the electrical disturbance which constitutes the wave has not yet reached this part of the path. The work along OT is equal and opposite to the work along WQ by symmetry. The net amount of work is equal to $TW \times H_y$. But this is equal to $1/c$ times the total current embraced by the path. In the present case the current is all of the displacement variety, so that it must be equal to the rate of increase of the flux of induction through the circuit, in the units we have been using. The

$$H_x' = 0, \quad H_y' = \frac{c}{a_2\mu_2}\cos\theta_2 X_2 e^{\iota\beta_2(a_2t - y\sin\theta_2 - z\cos\theta_2)},$$

$$H_z' = -\frac{c}{a_2\mu_2}\sin\theta_2 X_2 e^{\iota\beta_2(a_2t - y\sin\theta_2 - z\cos\theta_2)},$$

and the reflected wave by

$$E_x'' = X_3 e^{\iota\beta_3(a_3t - y\sin\theta_3 - z\cos\theta_3)}, \quad E_y'' = 0, \quad E_z'' = 0;$$

$$H_x'' = 0, \quad H_y'' = \frac{c}{a_3\mu_1}\cos\theta_3 X_3 e^{\iota\beta_3(a_3t - y\sin\theta_3 - z\cos\theta_3)},$$

$$H_z'' = -\frac{c}{a_3\mu_1}\sin\theta_3 X_3 e^{\iota\beta_3(a_3t - y\sin\theta_3 - z\cos\theta_3)}.$$

Since the boundary conditions have to be satisfied at all values of the time, the time factor in these expressions must be the same for all of them. Thus $\beta_1 a_1 = \beta_2 a_2 = \beta_3 a_3$. Also the velocity a is determined by the medium, so that $a_1 = a_3$. Hence $\beta_3 = \beta_1$ and $\beta_2 = \frac{a_1}{a_2}\beta_1$. The boundary conditions must be satisfied also for all values of y, so that the exponential factor in y must be common to each of the vectors. Thus $\beta_1 \sin\theta_1 = \beta_2 \sin\theta_2 = \beta_3 \sin\theta_3 = \beta_1 \sin\theta_3$. Therefore $\sin\theta_3 = \sin\theta_1$, so that the angle of reflexion must be equal in magnitude to the angle of incidence. Also

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{\beta_2}{\beta_1} = \frac{a_1}{a_2} \quad \text{......................(17)}.$$

Thus the refractive index or the ratio of the sine of the angle of incidence to the sine of the angle of reflexion is equal to the ratio of the velocities of propagation of the light in the two media. Also with the convention as to the signs of a_1 and a_3 and of $\sin\theta_1$ and $\sin\theta_3$ which is here adopted, in conformity with general usage, $\cos\theta_3 = -\cos\theta_1$.

The boundary conditions will now be satisfied if

$$X_1 + X_3 = X_2 \quad \text{....................(18)},$$

$$\frac{c}{a_1\mu_1}\cos\theta_1 (X_1 - X_3) = \frac{c}{a_2\mu_2}\cos\theta_2 X_2 \quad \text{.........(19)},$$

$$\frac{c}{a_1}\sin\theta_1 (X_1 + X_3) = \frac{c}{a_2}\sin\theta_2 X_2 \quad \text{............(20)}.$$

The expressions for the refracted and reflected waves are obtained by replacing the suffix 1 by the suffixes 2 and 3 respectively. As before, the exponential factors must be identical when $z=0$ since the boundary conditions are to hold for all values both of t and y. It follows that

$$a_3 = a_1,\ \sin\theta_3 = \sin\theta_1 \text{ and } a_1/a_2 = \beta_2/\beta_1 = \sin\theta_1/\sin\theta_2.$$

From the condition that the tangential magnetic force is continuous, we have

$$H_1 + H_3 = H_2.$$

The continuity of the tangential electric force gives

$$\frac{\cos\theta_1}{\kappa_1 a_1}(H_1 - H_3) = \frac{\cos\theta_2}{\kappa_2 a_2} H_2,$$

and the continuity of the normal electric induction

$$\frac{\sin\theta_1}{a_1}(H_1 + H_3) = \frac{\sin\theta_2}{a_2} H_2.$$

As before, the first and last of these equations are identical. For $\mu_1 = \mu_2 = 1$ we have $\kappa_1/\kappa_2 = a_2^2/a_1^2$ so that the second equation may be written

$$\sin\theta_1 \cos\theta_1 (H_1 - H_3) = \sin\theta_2 \cos\theta_2 . H_2.$$

Solving for H_2 and H_3 we find

$$H_2 = \frac{2\sin 2\theta_1}{\sin 2\theta_1 + \sin 2\theta_2} H_1 \quad\text{.................}(24)$$

and

$$H_3 = \frac{1 - \dfrac{\sin\theta_1\cos\theta_1}{\sin\theta_2\cos\theta_2}}{1 + \dfrac{\sin\theta_1\cos\theta_1}{\sin\theta_2\cos\theta_2}} H_1$$

$$= \frac{1 + \tan^2\theta_1 - \tan\theta_1(\cot\theta_2 + \tan\theta_2)}{1 + \tan^2\theta_1 + \tan\theta_1(\cot\theta_2 + \tan\theta_2)} H$$

$$= \frac{\tan(\theta_2 - \theta_1)}{\tan(\theta_2 + \theta_1)} H_1 \quad\text{.....................}(25).$$

When $\theta_1 + \theta_2 = \pi/2$, $\tan(\theta_1 + \theta_2) = \infty$; so that for the particular value of $\theta_1 = \theta_p = \pi/2 - \theta_2$, the intensity of the reflected wave is zero. This angle is known as the polarizing angle. When θ_1 exceeds θ_p, H_3 becomes negative so that there is a sudden change of phase at the polarizing angle.

Let us suppose that waves of frequency p impinge on the boundary of such a medium. We shall see that the boundary conditions of the problem can be satisfied by means of a transmitted and a reflected wave both of which have the same frequency p. The time variation of all the vectors can therefore be appropriately represented by the common factor $e^{\iota pt}$.

Let us first consider the propagation of a wave of this frequency inside the conducting medium. The electric and magnetic forces have to satisfy the equations

$$\text{rot } H = \frac{1}{c}\left(\sigma + \kappa \frac{\partial}{\partial t}\right) E,$$

and
$$\text{rot } E = -\frac{\mu}{c}\frac{\partial H}{\partial t}.$$

Since $\dfrac{\partial}{\partial t} \equiv \iota p$ for the waves under consideration, these equations are equivalent to

$$\text{rot } H = \frac{1}{c}\left(\frac{\sigma}{\iota p} + \kappa\right)\frac{\partial}{\partial t} E = \frac{\kappa_1}{c}\frac{\partial E}{\partial t}.$$

and
$$\text{rot } E = -\frac{\mu}{c}\frac{\partial H}{\partial t}.$$

Thus the results which we have obtained for a non-conducting medium will still hold if we replace the dielectric constant of the medium κ by the complex quantity $\kappa_1 = \kappa + \sigma/\iota p$. The components of the electric and magnetic vectors still satisfy the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \nabla^2 u,$$

but the constant a^2 is now complex and equal to $c^2/\mu\,(\kappa + \sigma/\iota p)$. Confining ourselves to the case in which the waves are propagated along the axis of z, since everything is proportional to $e^{\iota pt}$, the equation of propagation for this case reduces to

$$\frac{\mu}{c^2}(\iota p\sigma - p^2\kappa)\, u = \frac{\partial^2 u}{\partial z^2}.$$

The solution of this, appropriate to plane waves, is

$$u = e^{\pm(\alpha + \iota\beta)\, z},$$

where
$$(\alpha + \iota\beta)^2 = \frac{\mu}{c^2}(\iota p\sigma - p^2\kappa),$$

to κ in calculating α from (28) does not exert any important influence on the result. The formula also does not suggest the observed difference in colour between the incident and transmitted light. These discrepancies arise from the fact that the current is carried by discrete electrons, with the consequence that both σ and κ are functions of p. The reason for this will be made clearer in the sequel; we shall now consider the phenomena which attend the reflexion of light at a conducting surface.

Metallic Reflexion.

The problem of metallic reflexion is very similar to that furnished by the case of reflexion at the boundary between two insulating media. The same conditions as to continuity of the tangential electric and magnetic forces and of the normal electric and magnetic inductions have to be satisfied in both cases. The difference arises from the conducting power of the metallic medium, and we have seen that the type of theory proper to an insulating medium accounts for the propagation of waves in a conducting medium if we replace the dielectric constant by the complex quantity $\kappa_2' - \iota\sigma_2/p$, where σ_2 is the conductivity of the medium, κ_2' its dielectric coefficient and p the frequency of the waves. It is natural therefore to see if the boundary conditions cannot be fitted by the method previously adopted, the only change made being that the real quantity κ_2 in the former problem is replaced by the complex quantity $\kappa_2' - \iota\sigma_2/p$, where κ_2' is the real dielectric coefficient. We shall consider here only the case of waves polarized in the plane of incidence. A more complete discussion may be found in Drude's *Lehrbuch der Optik*, p. 334.

The incident wave being

$$E_x = X_1 e^{\iota\beta_1(a_1 t - y\sin\theta_1 - z\cos\theta_1)}, \qquad E_y = 0, \qquad E_z = 0;$$

$$H_x = 0, \qquad H_y = \frac{c}{a_1\mu_1}\cos\theta_1 X_1 e^{\iota\beta_1(a_1 t - y\sin\theta_1 - z\cos\theta_1)},$$

$$H_z = -\frac{c}{a_1\mu_1}\sin\theta_1 X_1 e^{\iota\beta_1(a_1 t - y\sin\theta_1 - z\cos\theta_1)},$$

the refracted and reflected waves will be given by similar expressions with the suffix 1 replaced by the suffixes 2 and 3

The following table gives the values of the product $(100-R)\sigma^{\frac{1}{2}}$, obtained by this method for the residual rays of wave-length $25\cdot5\times10^{-4}$ cm. which remain after successive reflexions from surfaces of fluorite. Considering the smallness of $100-R$ the agreement of the experimental and computed values is remarkably good.

Metals	Emission-power $I=(100-R)$ for $\lambda=25\cdot5\times10^{-4}$ cm. and 170°		Product $(100-R)\sqrt{\sigma_{170}}$ taken from column 2	
	Computed by Formula (36)	Observed		
Silver	1·15	1·13	7·07	Mean 7·33
Copper	1·27	1·17	6·67	
Gold	1·39	1·56	8·10	
Aluminium	1·60	1·97	8·91	
Zinc	2·27	2·27	7·24	
Cadmium	2·53	2·55	7·29	
Platinum	2·96	2·82	6·88	
Nickel	3·16	3·20	7·33	
Tin	3·23	3·27	7·32	
Steel	3·99	3·66	6·62	
Mercury	7·55	7·66	7·33	
Bismuth	10·09	(25·6)	(18·3)	
"Rotgiss"	2·73	2·70	7·16	Mean 7·41
Manganin	4·69	4·63	7·16	
Constantan	5·05	5·20	7·43	
Patent Nickel *P*	3·77	4·05	7·77	
Patent Nickel *M*	4·28	4·45	7·53	

There are two results of this investigation which are of special interest. In the visible part of the spectrum it is known that the reflecting power of metals does not agree with the predictions of Maxwell's Theory in its simple form, so that the present experiments determine the boundary of the region where other considerations have to be taken into account. The experimental results are in accordance with the simple theory when the wave-length of the radiation is equal to $25\cdot5\times10^{-4}$ cm.

The other point relates to the magnetic qualities of metals. The computed values in the tables have been obtained by putting $\mu_1=\mu_2=1$ and $\kappa_1=\kappa_2'=1$ for all the metals. We see that the agreement in the case of the magnetic metals and alloys such as nickel, steel and so on is just as good as in the case of the non-magnetic metals. It follows that for oscillations of the frequency of those experimented with, the magnetic metals behave as though they were non-magnetic.

suppose that under the action of an electric intensity E the different electrons in the molecule denoted by the suffixes $1, 2, \ldots n$, undergo displacements $x_1, x_2, \ldots x_n$ from the equilibrium configuration. These displacements are supposed to be all in the same direction, which is that of the applied intensity E. We shall suppose that when a displacement x_1, for example, takes place, there is a force of restitution called into play which is proportional and parallel to x_1, and independent of $x_2, x_3, \ldots x_n$. This assumption that the force of restitution, which we shall denote by x_1/λ_1, is independent of $x_2, x_3, \ldots x_n$, can only be justified as a rough approximation, since it is clear that each of the displacements $x_2, \ldots x_n$ is equivalent to a doublet of moment $e_2x_2, e_3x_3, \ldots e_nx_n$, where the e's are the charges of the respective electrons, and it is evident that each of these doublets will give rise to a force on the electron e_1 proportional to its moment. The case in which this approximation is not made will be considered later (p. 169).

In addition to the forces of restitution each of the displaced electrons may, in general, be acted on by one or more forces of each of the following types:

(1) The external impressed force. This arises from the external electric intensity E, and it might be thought at first sight that its magnitude would be given by $Ee_1, Ee_2, \ldots Ee_n$. The case here is, however, identical with that discussed at the end of Chapter IV, and the argument pursued there shows that the value of the actual force of external origin acting on the electrons is given by $(E + aP)e_1 \ldots (E + aP)e_n$, where P is the polarization of the medium, and a is a constant, which, if the medium is sufficiently symmetrical, is equal to one-third.

(2) There will also be forces of frictional type tending to oppose the motion. These forces may be represented by a term proportional to the velocity, which in the case of the sth electron for example we may denote by $-\beta_s\dot{x}_s$ where β_s is a constant.

The precise mechanism of these forces is not yet properly understood. We shall see later that when an electrified particle is accelerated it emits radiation, and the emission of this radiation gives rise to a reaction $\frac{e^2}{6\pi c^3}\dddot{x}$ (see p. 266). For a simple harmonic vibration proportional to $e^{\iota pt}$ this is equivalent to a force energy of the resonating system will be converted into energy of other types, and that particular degree of freedom will possess comparatively little energy when the original system is reformed. The general results of this view are in qualitative agreement with those which follow from the assumption of a retarding force proportional to $\dot{x}$. They show the same gradual change in phase in passing through an absorption band, and indicate a maximum absorption and a refractive index equal to unity, when the period of the light is approximately coincident with the natural period of the substance.

In view of the facts that absorption does occur, and that its mode of occurrence is still doubtful, we shall for the present content ourselves with the assumption of a term, in the expression for the force, proportional to $-\dot{x}$. This term is to be regarded, not as the expression of a fundamental truth, but as a simple and convenient mathematical approximation whose consequences simulate the observed effects.

(3) In general we shall have to take account of the effect of the presence of a magnetic field. We have seen (Chap. VI, p. 68) that a magnetic field of intensity H acting on a moving electron, charge e_s, gives rise to a force whose components are

$$\frac{e_s}{c}(H_z \dot{y}_s - H_y \dot{z}_s),$$

$$\frac{e_s}{c}(H_x \dot{z}_s - H_z \dot{x}_s),$$

$$\frac{e_s}{c}(H_y \dot{x}_s - H_x \dot{y}_s).$$

The total force acting on the electron will be compounded of all the forces mentioned, and, in the most general case, its equations of motion will be

$$\begin{aligned} m_s \frac{\partial^2 x_s}{\partial t^2} &= e_s(E_x + \alpha_s P_x) - \frac{x_s}{\lambda_s} - \beta_s \frac{\partial x_s}{\partial t} + \frac{e_s}{c}\left(H_z \frac{\partial y_s}{\partial t} - H_y \frac{\partial z_s}{\partial t}\right), \\ m_s \frac{\partial^2 y_s}{\partial t^2} &= e_s(E_y + \alpha_s P_y) - \frac{y_s}{\lambda_s} - \beta_s \frac{\partial y_s}{\partial t} + \frac{e_s}{c}\left(H_x \frac{\partial z_s}{\partial t} - H_z \frac{\partial x_s}{\partial t}\right), \\ m_s \frac{\partial^2 z_s}{\partial t^2} &= e_s(E_z + \alpha_s P_z) - \frac{z_s}{\lambda_s} - \beta_s \frac{\partial z_s}{\partial t} + \frac{e_s}{c}\left(H_y \frac{\partial x_s}{\partial t} - H_x \frac{\partial y_s}{\partial t}\right) \end{aligned} \quad \ldots\ldots(1).$$

are so distinctive that the magnitude of the molecular refractivity is regarded as a valuable aid in determining the constitution of organic compounds. On the whole, the general constancy of the atomic refractivities, and the definiteness of the changes when changes do take place, supply us with striking evidence that the structure of the atoms is not very profoundly disturbed by chemical combination.

Dispersion.

We shall now consider the way in which, according to our formulae, the refractive index depends upon the frequency or wave-length of the incident vibrations. Formula (5) may be written somewhat more simply, in the form

$$\frac{m^2-1}{m^2+a}=\frac{1}{a+1}\sum_1^n \frac{\nu_s e_s^2}{m_s(p_s^2-p^2)},$$

where $a=\frac{1-\alpha}{\alpha}$. Now since the incident vibrations vary as $e^{\iota pt}$, and the natural vibrations as $e^{\iota p_1 t}$, $e^{\iota p_2 t}$, etc., they are respectively periodic in times $2\pi/p$, $2\pi/p_1$, $2\pi/p_2$, etc. If λ, λ_1, λ_2, etc., denote the corresponding wave-lengths measured in the free aether (not in the substance traversed by the waves)

$$\lambda=2\pi c/p, \qquad \lambda_1=2\pi c/p_1, \qquad \lambda_2=2\pi c/p_2, \text{ etc.}$$

Thus
$$\frac{m^2-1}{m^2+a}=\frac{1}{4\pi^2c^2(1+a)}\sum_1^n \frac{\nu_s e_s^2\lambda^2\lambda_s^2}{m_s(\lambda^2-\lambda_s^2)}.$$

But
$$\frac{\lambda^2\lambda_s^2}{\lambda^2-\lambda_s^2}=\lambda_s^2+\frac{\lambda_s^4}{\lambda^2-\lambda_s^2};$$

so that
$$\frac{m^2-1}{m^2+a}=\frac{1}{4\pi^2c^2(1+a)}\left\{\sum_1^n \frac{\nu_s e_s^2\lambda_s^2}{m_s}+\sum_1^n \frac{\nu_s e_s^2\lambda_s^4}{m_s(\lambda^2-\lambda_s^2)}\right\}.$$

Now when λ is infinite $\sum_1^n \frac{\nu_s e_s^2\lambda_s^4}{m_s(\lambda^2-\lambda_s^2)}=0$, since every term in the summation vanishes. But when the wave-length is infinitely long and the period is infinitely slow, the case under consideration approaches continuously to that of an electrostatic field. The quantity κ, which enters into the equations of propagation, must therefore become identical with the dielectric coefficient as measured by electrostatic methods. Hence for infinitely long waves $m^2=\kappa$, the dielectric coefficient of the medium. We therefore have

$$\frac{\kappa-1}{\kappa+a}=\frac{1}{4\pi^2c^2(1+a)}\sum_1^n \frac{\nu_s e_s^2\lambda_s^2}{m_s} \qquad \text{...............(7)},$$

and
$$\frac{m^2-1}{m^2+a}=\frac{\kappa-1}{\kappa+a}+\frac{1}{4\pi^2c^2(1+a)}\sum_1^n\frac{\nu_s e_s^2\lambda_s^4}{m_s(\lambda^2-\lambda_s^2)}\quad\ldots\ldots(8).$$

ν_s, e_s, λ_s and m_s are constants characteristic of the material under consideration, so that (8) may be written in the form

$$\frac{m^2-1}{m^2+a}=\frac{\kappa-1}{\kappa+a}+\frac{C_1}{\lambda^2-\lambda_1^2}+\frac{C_2}{\lambda^2-\lambda_2^2}+\ldots+\frac{C_n}{\lambda^2-\lambda_n^2}\quad\ldots(9),$$

where C_1, C_2, … C_n are constant quantities characteristic of the substance. There will be one of the terms C_1, C_2, etc., corresponding to each of the natural periods.

In the case of transparent colourless substances the natural periods must be either in the infra-red or in the ultra-violet part of the spectrum. If they are in the infra-red the synchronous wave-lengths λ_r will be large compared with λ, and if they are in the ultra-violet the λ_v's will be small compared with λ. Thus, as an approximation, we may write (9) in the form

$$\frac{m^2-1}{m^2+a}=\frac{\kappa-1}{\kappa+a}-\Sigma\frac{C_r}{\lambda_r^2}+\Sigma\frac{C_v}{\lambda^2}$$

$$=\text{Constant}+\Sigma\frac{C_v}{\lambda^2},\ \text{approx}\ldots\ldots\ldots\ldots(9a).$$

In these formulae we have neglected the fractions $\frac{\lambda_v^2}{\lambda^2}$ and $\frac{\lambda^2}{\lambda_r^2}$ compared with unity. Since the refractive index m changes roughly in the same way as $\frac{m^2-1}{m^2+a}$, it follows from (9a) that for these substances the refractive index will increase continuously as λ diminishes, in this region. Since the transparent colourless substances were the first to have their dispersion investigated, this type of dispersion is said to be normal. As we shall see, it can only be said to be normal provided we are a long way from the natural frequencies of the substances.

The behaviour of formulae (8) and (9) in the case of light whose frequency is close to that of the natural vibrations of the substance is most interesting. When $p=p_s$ for example, then $\lambda=\lambda_s$, and the corresponding term in (9) becomes infinite. If λ is slightly less than λ_s, then $C_s/(\lambda^2-\lambda_s^2)$ has a large negative value; and if λ is slightly greater than λ_s, then $C_s/(\lambda^2-\lambda_s^2)$ has a large positive value. As λ approaches λ_s from smaller values of which connect the refractive index with the wave-length or frequency of the transmitted light, have been in vogue from time to time. Cauchy's formula

$$m = A_1 + A_2\lambda^{-2} + A_3\lambda^{-4} + \ldots$$

was the first to have a theoretical justification. Cauchy developed it from general considerations of wave theory on the hypothesis that the distance between the vibrating particles of the medium could not be regarded as completely negligible when compared with the wave-length. The formula thus obtained gives a fairly satisfactory representation, when three terms are used, of the dispersion of a number of substances. As it makes the refractive index increase continuously with diminishing wave-lengths (the constants are all positive), it fails absolutely to account for anomalous dispersion.

The idea that time rather than length was the determining common factor of the light and the matter which gave rise to dispersion, appears to have occurred first to Maxwell*. Maxwell supposed that when the atoms (we should now say electrons) were displaced, forces of restitution were called into play, and that there was also a resistance to their motion. A similar idea occurred somewhat later to Sellmeier, after whom the formula to which this theory gives rise is usually named. In the case of a single mode of vibration and in the absence of friction, the case contemplated by Sellmeier, the formula is

$$m^2 = A + \frac{B}{\lambda^2 - \lambda_s^2},$$

where λ is the wave-length of the light, and λ_s that corresponding to the natural vibration of the substance. In the case in which a (p. 73) = 0, formula (5) becomes

$$\begin{aligned} m^2 &= 1 + \sum_1^n \frac{\nu_s e_s^2}{m_s(p_s^2 - p^2)} \\ &= 1 + \sum_1^n \frac{\nu_s e_s^2 \lambda_s^2}{4\pi^2 c^2 m_s^2} + \sum_1^n \frac{\nu_s e_s^2 \lambda_s^4}{4\pi^2 c^2 m_s^2 (\lambda^2 - \lambda_s^2)} \quad \ldots\ldots(10), \end{aligned}$$

so that Sellmeier's formula can be regarded as the particular case of formula (5) which arises when we put $a = 0$ and $n = 1$.

It is still an open question whether the best dispersion formula is the Lorentz formula (5), or a generalized Sellmeier formula (see p. 176). It is however pretty clear that formula (10) is unsatisfactory, since, other things being equal, it makes m^2-1 vary as the density of the medium, and this is a great deal further from the truth than $\frac{m^2-1}{m^2+a} \propto \rho$. Both (5) and (10) give the same general kind of behaviour of m^2 when any one of the zeros $\lambda=\lambda_1$, $\lambda=\lambda_2$, etc., is crossed. The infinities of m^2 are however in different positions in the two cases. Thus in a formula of type (10), viz.

---

* Lord Rayleigh's *Coll. Papers*, vol. IV. p. 413.

$$m^2 = A_0 + \frac{A_1}{\lambda^2-\lambda_1^2} + \frac{A_2}{\lambda^2-\lambda_2^2} + \dots,$$

in the neighbourhood of, let us say $\lambda=\lambda_s$, we may put

$$m^2 = q_s + \frac{A_s}{\lambda^2-\lambda_s^2},$$

the rate of variation of the terms not involving λ_s being comparatively negligible. We observe that m^2 is negative from $\lambda=\lambda_s$ to $\lambda=\sqrt{\lambda_s^2-\frac{A_s}{q_s}}=\lambda_s\left(1-\frac{1}{2}\frac{A_s}{\lambda_s^2 q_s}\right)$ approx. Thus there is a range of wave-length $\delta\lambda_s = A_s/2q_s\lambda_s$ for which m is imaginary. In the case of a formula of type (5) we may write in the neighbourhood of an absorption band

$$\frac{m^2-1}{m^2+a} = q + \frac{A}{\lambda^2-\lambda_s^2},$$

or

$$m^2 = \frac{(1+aq)(\lambda^2-\lambda_s^2)+Aa}{(1-q)(\lambda^2-\lambda_s^2)-A}$$

$$= \frac{1+aq}{1-q} + A\frac{1+a}{(1-q)^2}\Big/(\lambda^2-\lambda_s'^2) \dots\dots\dots\dots\dots (11),$$

where

$$\lambda_s'^2 = \lambda_s^2 + A/(1-q).$$

It is clear from (11) that m^2 changes from $+\infty$ to $-\infty$ as λ crosses from $\lambda>\lambda_s'$ to $\lambda<\lambda_s'$. m^2 then becomes continuously smaller numerically, but remains negative as to sign until it reaches the value zero, which is determined by

$$1 + aq + \frac{A(1+a)}{(1-q)(\lambda^2-\lambda_s'^2)} = 0,$$

or $\lambda^2 = \lambda_s^2 - Aa/(1+aq)$. Thus m is imaginary from

$$\lambda = \{\lambda_s^2 + A/(1-q)\}^{\frac{1}{2}} \text{ to } \lambda = \{\lambda_s^2 - Aa/(1+aq)\}^{\frac{1}{2}}.$$

Many of the aniline dyes exhibit phenomena of this character in the visible spectrum. In fact something similar arises whenever the medium possesses intense absorption; so that it will be more convenient to consider the case of the aniline dyes when the theory of absorbing media has been discussed. It is clear from the considerations which have been urged that it is not *necessary* to have absorption of the type which is accompanied by degradation of energy for the rays to be unable to traverse the medium. It is only necessary that the period of the light vibrations should agree with one of the natural periods of the medium.

The best examples of this type of phenomenon have been found in the behaviour, in the infra-red region of the spectrum, of a number of insulators which are quite transparent to light in the visible spectrum. The most conspicuous examples are quartz, rock-salt, sylvite and fluorite. If the reflecting power of quartz, for example, is examined, it is found to be small, for normal incidence, for all wave-lengths from the visible spectrum up to about

$$7{\cdot}6\,\mu\,(1\mu = 10^{-4}\text{ cm.}).$$

It then begins to increase rapidly as λ increases, until at a wave-length in the neighbourhood of $8{\cdot}1\,\mu$ quartz is almost as good a reflector of radiation as a metal. This state of things continues up to about $9\,\mu$, when the reflecting power begins to diminish. The transparency varies in the opposite way to the reflecting power. In fact between 8·1 and $9\,\mu$ quartz is so opaque that Nichols* was unable to detect any radiation through a layer of it only 2·5 wave-lengths in thickness. The relation between the reflecting power and the transparency of quartz is exhibited in the accompanying diagram (Fig. 25), which represents the results of Nichols's experiments.

Since the rays which correspond to the natural periods of substances are incapable of entering them, and so are always almost totally reflected, we are furnished with a new means of investigating the optical periods of substances. This method, due to Rubens, consists in submitting a beam of radiation from some source, such as a Nernst glower, to a series of successive reflexions, at incidences as nearly normal as possible, from surfaces of the material under investigation. The rays which are obtained after

---

* *Ann. der Phys.* vol. LX. p. 401 (1897).

Thus
$$2n^2 = \sqrt{A^2 + \frac{B^2 + 2AB\gamma}{\gamma^2 + \delta^2}} + A + \frac{B\gamma}{\gamma^2 + \delta^2} \quad \text{.........(21)},$$

$$2n^2\kappa^2 = \sqrt{A^2 + \frac{B^2 + 2AB\gamma}{\gamma^2 + \delta^2}} - A - \frac{B\gamma}{\gamma^2 + \delta^2} \quad \text{.........(22)}.$$

Feeble Absorption.

There is one particular case, which is of importance in nature, in which a considerable simplification of these formulae may be effected. A large number of substances exhibit absorption in varying degree throughout their volume, and are said to possess body colour. This is the case with most coloured solutions and minerals, and in fact with most coloured substances which are transparent to a considerable proportion of the visible spectrum. In these cases the absorption is small in a distance comparable with the wave-length of light, although it may be considerable in a distance comparable with 1 centimetre.

Consider the behaviour of waves for which p has a value such as to make $\gamma = 0$. We see from p. 165 that the absorption causes the intensity of the light to fall off as $e^{-2\pi\frac{n\kappa}{\lambda}z}$. In a distance $z = \frac{\lambda}{2\pi}$ it will therefore diminish in the proportion of $e^{-n\kappa}$ to unity. Since the absorption in distances of this order is very small it follows that $e^{-n\kappa}$ must be very nearly equal to unity; so that $n\kappa$ must be a very small quantity. Turning to formula (22) we see that $n\kappa$ will only be small when $\gamma = 0$, provided that B/δ is a small quantity. It follows that δ is a large quantity compared with B, and also that it is large compared with γ within a reasonable distance of the value of p corresponding to $\gamma = 0$. In this region we may use, as an approximation, the formula

$$n\kappa = \frac{B\delta}{2A^{\frac{1}{2}}(\gamma^2 + \delta^2)} \quad \text{....................(23)}.$$

At the frequency for which the absorption is a maximum $\frac{\partial}{\partial p}(n\kappa)$ and therefore $\frac{\partial}{\partial p}\left(\frac{\delta}{\gamma^2 + \delta^2}\right)$ vanishes. Hence

$$mp_0^2 - \frac{ave^2}{1 - aq} - mp^2 = \frac{\beta^2 p^2}{mp_0^2 - \frac{ave^2}{1 - aq} + 3mp^2}.$$

Generalized Theory of Dispersion.

In order properly to realize the limitations and approximations to which the foregoing theory of dispersion is subject, it is desirable to consider it briefly from a somewhat wider standpoint. From what has been said it is clear* that dispersion is essentially a dynamical problem in which the machinery is determined by the fundamental structure of the atom. Unfortunately we know very little about this structure, so that the necessary data cannot be stated very explicitly. Fortunately the methods of generalized dynamics enable us to find out a good deal about the behaviour of such a system even when we do not know much about its exact constitution.

In order to determine the motion it is necessary that certain functions of the state of the system should be known or, at least, be capable of definite specification. These are the Kinetic Energy, the Potential Energy, the Dissipation Function and the function which is equal to the work of the external forces. Let us consider these briefly in order.

Regarded as a dynamical system the optical medium consists of a system of electrons which may be treated as point charges subject, when undisturbed, to unknown conditions of equilibrium. The equilibrium is not necessarily a static one but may involve motion in orbits.

In any event the expression for the kinetic energy is quite simple, since it is equal to the sum of the energies of the individual electrons. If there are n electrons in any sufficiently large element of volume of the medium the kinetic energy δT which belongs to this element is

$$\delta T = \frac{1}{2} \sum_{s=1}^{s=n} m_s (\dot{x}_s^2 + \dot{y}_s^2 + \dot{z}_s^2) \quad \ldots\ldots\ldots\ldots\ldots(28),$$

where m_s is the mass and $\dot{x}_s$, $\dot{y}_s$ and $\dot{z}_s$ denote the components of the velocity of the sth electron.

To obtain an expression for the potential energy is much more complicated. If the equilibrium state involves steady motion, the

---

* This is only true provided dynamics is adequate completely to account for the behaviour of atoms in this respect. This point is now doubtful, but, at least, it is of interest to examine the results to which generalized dynamics leads.

The potential energy of e_s in its undisplaced position, due to the doublet arising from e_a the moment of which is $e_a x_a$, is

$$e_s e_a x_a \frac{\partial}{\partial x_a}\left(\frac{1}{r}\right) = -e_s e_a x_a \frac{x}{r^3}.$$

The corresponding quantity for the displaced position of e_s is

$$e_s e_a x_a \frac{\partial}{\partial x_a}\left(\frac{1}{r'}\right) = -e_s e_a x_a \frac{x + x_s}{r'^3}.$$

There will therefore be a term in the potential energy of e_s, due to any other electron, which is of the form $e_s e_a x_s x_a f(r, \theta)$ where $f(r, \theta)$ depends upon the undisplaced positions of the electrons in space.

Similar remarks apply to the y and z components of the displacements of the electrons, so that the complete expression for the potential energy of the element of volume may be written

$$\delta W = \sum_{r=1}^{r=n} \sum_{s=1}^{s=n} \{A_{rs} x_r x_s + B_{rs} y_r y_s + C_{rs} z_r z_s + F_{rs} x_r y_s + G_{rs} x_r z_s + H_{rs} y_r z_s\} \ldots (29),$$

where the coefficients A, B, C, F, G, H involve the fundamental structure of the medium, but are independent of the displacements $x_r, y_r \ldots z_s$.

The nature of this expression for the potential energy calls for a little fuller consideration. The cross coefficients A_{rs} etc. are proportional to the inverse cube of the mutual distance of the electrons involved, so that they are small except for pairs of electrons which are quite near one another. On the other hand the number of electrons at a distance between r and $r + dr$ varies as $r^2 dr$; but, on account of the periodic character of the phenomena, the distant electrons are in layers which exert opposite and approximately equal effects. Thus the contribution to the potential energy of a particular electron, which arises from the cross terms, will come almost entirely from other electrons in its immediate neighbourhood; in other words the whole of this potential energy may be considered to arise from local causes. It is therefore legitimate to express it as a summation over the element of volume if this is taken fairly large. Unfortunately the size of the appropriate element will depend to some extent on the period of the vibrations, so this process can only be regarded as an approximation after all.

where the mechanical force on an electric charge whose velocity is V relative to the measuring system is, per unit charge,

$$F = E + \frac{1}{c}[VH] \qquad \text{.......................(5).}$$

It is necessary to show that these equations, which are associated with the name of Lorentz, are not inconsistent with any of our previous results. Looked at superficially they do appear to be inconsistent; since, by simply writing average values in equations (1) to (5) we do not arrive at equations which are obviously identical with those which we found to comprise the behaviour of dielectric and magnetic media in Chaps. III, V and VI. It is to be remembered, however, that the vectors defined as the electric and magnetic intensities and inductions respectively, in those chapters, were all average values of the true electric and magnetic intensities but formed in different ways. When this difference is taken into account the discrepancy will be seen to disappear. We shall now consider the equations in order from this point of view.

Equation (1) is supposed to apply to any element of volume however small. The corresponding equation div $D = \rho$ is an equation between average values, and only applies to an element of volume which contains a very large number of electrons. In order to compare them let us integrate (1) over any sufficiently large volume. We have

$$\iiint \rho \, d\tau = \iiint \left\{ \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right\} dx\,dy\,dz = \iint E_n \, dS,$$

where E_n represents the normal component of E at any point of the surface. But we have seen that the induction D_n is the average value of the force in the flat cavity perpendicular to D_n, so that the normal induction is nothing else than the average value of the intensity E taken over a surface perpendicular to it. This identification is only strictly true provided the surface is so large that the excess of polarization charges of a given sign inside of it is negligible. In other words, of the doublets whose axes are cut in two by the surface the difference between the number which leave their positive and those which leave their negative ends inside must be negligible. On the other hand, if material medium. If the strength of the current is i, (5) gives for the resultant force on the circuit the value $\frac{1}{c}\int[ids\,.\,H]$ taken round the circuit s. By an argument similar to that in Chap. v, p. 83, it follows that the force is equal to i/c times the rate of change of $\iint H_n dS$ over any surface having the same contour. The H_n in this integral is, of course, the normal component of the universal magnetic intensity. But, as we have seen,

$$\iint H_n dS = \iint B_n dS.$$

Thus it follows that (5) and ($5a$) are consistent and, incidentally, that the average value of H taken along a line normal to the direction of H is $\mu\bar{H} = B$. From this the analogy between the electric and magnetic vectors would lead us to expect that $\bar{E}_n = \frac{1}{s}\int E_n ds = \kappa\bar{E} = D$, the suffix n denoting that the vector is perpendicular to the direction of integration.

By dividing the space up by means of tubes of induction, it is clear that the average values of the universal expressions $\frac{1}{2}E^2$ and $\frac{1}{2}H^2$ for the electric and magnetic energy densities respectively are equal to $\frac{1}{2}\kappa\bar{E}^2$ and $\frac{1}{2}\mu\bar{H}^2$. This is only true provided we neglect constant terms which may be regarded as representing the intrinsic energy of the electrons and of the molecular magnets.

The Differential Equations satisfied by the Vectors when Charges are present.

In Chapter VII we were concerned with the solution of equations (1) to (4), and the extensions of them, which have just been considered, in the cases in which the density ρ of the charges was everywhere zero. The results thus obtained naturally applied to the propagation of electromagnetic effects in insulators, including the free aether as a particular case. We shall now consider the nature of the solutions in the more general case, when electric charges are present and contribute to the resulting phenomena by their motions and the forces they exert.

Let us now return to the consideration of the function F. This may be any function of the argument $t + r/c$. We shall suppose that it is such a function that it takes the value zero for all values of the argument except those in the immediate neighbourhood of the particular value zero. We shall then have $F(x) = 0$ unless x lies between, let us say, $\pm\epsilon$ where ϵ is a very small quantity. We shall also suppose that

$$\int_{-\epsilon}^{+\epsilon} F(x)\,dx = 1 \quad\text{......................(19).}$$

Since $F(x)$ is zero unless x is between $\pm\epsilon$ we evidently also have

$$\int_{-\infty}^{\infty} F(x)\,dx = \int_{-\epsilon}^{+\epsilon} F(x)\,dx = 1.$$

If the value of r is fixed

$$\int_{t_1}^{t_2} F\left(t + \frac{r}{c}\right) dt = \int_{t_1 + r/c}^{t_2 + r/c} F(x)\,dx = 1$$

provided $t_2 + r/c > \epsilon$ and $t_1 + r/c < -\epsilon$. Moreover, if we make ϵ infinitesimal but still suppose F to have the property $\int_{-\epsilon}^{+\epsilon} F(x)\,dx = 1$ we shall have, if ω is any function of r and t,

$$\int_{t_1}^{t_2} \omega F\left(t + \frac{r}{c}\right) dt = \omega' \int_{t_1}^{t_2} F\left(t + \frac{r}{c}\right) dt = \omega' \quad\text{.........(20),}$$

where ω' is the value of ω at the instant $t = -r/c$. This follows since, except when t lies between $\pm\epsilon$, $F = 0$, and throughout this infinitesimal interval ω may be considered constant.

Now let t_2 have a definite positive value and t_1 a very large definite negative value, $-t_1$ being so large that for all points in the enclosed volume $t_1 + r/c < -\epsilon$. Then the values of χ which occur in the integrated part of (18) all vanish. So also do the values of $\frac{\partial \chi}{\partial t}$, since the derivatives of χ are also zero except between $\pm\epsilon$.

We may write the term containing ω in (18) in the form

$$-\iiint \frac{1}{r}\,d\tau \int_{t_1}^{t_2} \omega F\left(t + \frac{r}{c}\right) dt.$$

This is equal to $-\iiint \frac{\omega'}{r}\,d\tau$ where ω' has the same meaning as in (20). In a similar way

$$\int_{t_1}^{t_2} dt \iint \chi \frac{\partial \psi}{\partial n}\,dS = \iint \frac{1}{r}\left(\frac{\partial \psi}{\partial n}\right)' dS,$$

at a time r/c earlier than the instant considered. This time r/c is equal to the time which would be required for light to travel from the distant element to the point P. The nature of the field is therefore such as would arise if each portion of it were constantly emitting *disturbances* which were *propagated from it in all directions with the velocity of light.*

When we come to the actual calculation of the values of E and H in particular cases it is found that equations (7) and (8) are unsuitable owing to the values of ω, given by the right-hand sides of them, being somewhat complicated. The calculations may be simplified by the introduction of two new functions, the scalar potential ϕ and the vector potential U, from which E and H may afterwards be derived by appropriate operations. We shall now prove that

$$H = \text{rot } U \quad \text{.........................(23)},$$

and

$$E = -\frac{1}{c}\frac{\partial U}{\partial t} - \text{grad }\phi \quad \text{.....................(24)}$$

if ϕ and U satisfy the equations

$$\nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = -\rho \quad \text{.......................(25)},$$

and

$$\nabla^2 U - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = -\frac{1}{c}\rho V \quad \text{.................(26)}.$$

We shall prove first of all that a function U always exists such that $H = \text{rot } U$. This function is in fact

$$U = \frac{1}{4\pi}\iiint \frac{\text{rot } H}{r}\, d\tau \quad \text{.................(27)}.$$

For if (27) is true we have

$$U_x = \frac{1}{4\pi}\iiint \frac{(\text{rot } H)_x}{r}\, d\tau = \frac{1}{4\pi}\iiint \frac{1}{r}\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right) d\tau.$$

Now the values of H in the integral refer to the different points of integration and not to the point at which U is measured. Let x, y, z be the coordinates of the point at which U is required and a, b, c the coordinates of the element of volume $d\tau$. Then $r^2 = (x-a)^2 + (y-b)^2 + (z-c)^2$ and the equation above may be written more clearly as

$$U_x = \frac{1}{4\pi}\iiint \left(\frac{\partial H_c}{\partial b} - \frac{\partial H_b}{\partial c}\right) \frac{da\,db\,dc}{\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}},$$

This equation is satisfied if $E+\frac{1}{c}\frac{\partial U}{\partial t}$ is the gradient of some scalar quantity $-\phi$. So that we may put

$$E=-\frac{1}{c}\frac{\partial U}{\partial t}-\text{grad}\,\phi.$$

This is the same as equation (24). It will be observed, however, that U and ϕ are not completely determined by the considerations which have been brought forward. The only condition we have imposed on U except (24) is that it should satisfy the equation $H=\text{rot}\,U$. Also ϕ may be any scalar quantity. If U_0 and ϕ_0 are particular values of U and ϕ which satisfy the equations under consideration, they will also be satisfied by

$$U=U_0-\text{grad}\,\psi \quad\text{and}\quad \phi=\phi_0+\frac{1}{c}\frac{\partial\psi}{\partial t} \quad\text{.........(29)},$$

where ψ is ~~some~~ any scalar function. We shall determine ψ by making it satisfy the condition

$$\text{div}\,U=-\frac{1}{c}\frac{\partial\phi}{\partial t} \quad\text{......................(30)}.$$

It is necessary to show that this condition can always be satisfied. Substituting the values (29) in equation (30) we get

$$\nabla^2\psi-\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2}=\text{div}\,U_0+\frac{1}{c}\frac{\partial\phi_0}{\partial t}.$$

There is always some value ψ which will satisfy this equation, so that (30) can always be satisfied.

We have

$$\rho=\text{div}\,E=-\frac{1}{c}\frac{\partial}{\partial t}(\text{div}\,U)-\text{div}\,\text{grad}\,\phi,$$

whence making use of (30)

$$\nabla^2\phi-\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}=-\rho.$$

We also have

$$\text{rot}\,H=\frac{1}{c}\left(\frac{\partial E}{\partial t}+\rho V\right).$$

Substituting from (23) and (24)

$$\begin{aligned}\text{rot}\,\text{rot}\,U&=\text{grad}\,\text{div}\,U-\nabla^2 U\\&=\frac{1}{c}\left\{-\frac{1}{c}\frac{\partial^2 U}{\partial t^2}-\text{grad}\frac{\partial\phi}{\partial t}+\rho V\right\},\end{aligned}$$

and

$$\operatorname{div} U = -\frac{1}{c}\frac{\partial \phi}{\partial t},$$

so that

$$\nabla^2 U - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = -\frac{1}{c}\rho V.$$

Thus (25) and (26) are the equations satisfied by the scalar and vector potentials respectively. Conversely if ϕ and U satisfy (25) and (26) the vectors E and H will be given by (23) and (24).

In the light of our discussion of equations (7) and (8) it is clear that the values of the potentials are

$$\phi = +\frac{1}{4\pi}\iiint \frac{\rho'}{r}\, d\tau$$

and

$$U = \frac{1}{4\pi c}\iiint \frac{(\rho V)'}{r}\, d\tau,$$

where the dashes denote that in carrying out the integrations the values of ρ and ρV, respectively, at the instant $t = -r/c$, previous to that for which the integrals are being evaluated, have to be substituted.

Electron at Rest and in Uniform Motion.

As an illustration of the results which we have just obtained we shall consider the case of a single electron. If the electron has always been at rest then V is always zero; so that the vector potential U vanishes. Moreover ρ' becomes identical with ρ for every point, since the position at any previous instant is the same as the instantaneous position. Thus the scalar potential is identical with the ordinary static potential, the electric intensity is identical with the usual value of electrostatics and the magnetic force vanishes. The solution in this case is identical with the results of the usual electrostatic theory.

Next consider an electron which is moving and has always moved with a uniform velocity w in a straight line parallel to the axis of z. Consider the values of the two potentials at any point P_1 at an instant t_1. They will not be determined by the instantaneous state of the electron, but by its state at some previous instant t_1'. t_1' will in fact be given by the equation

$$t_1' = t_1 - r_1'/c,$$

Now $\frac{1}{2}A^2$ is the mean energy present in unit volume of the wave, so that Poynting's Theorem represents this energy as flowing along with the velocity c. Since this is in accordance with the results of observation the application of Poynting's Theorem to questions relating to radiation evidently rests on very solid grounds.

Forces exerted on the Charges.

In the last section but one we have considered the rate of working or "activity" of the forces acting on an enclosed electrical system. We shall now consider the value of the resultant force acting on a similar system. On the electron theory of matter the results will be applicable to any material system since, on this theory, the force acting on a material system is the aggregate effect of the electric and magnetic forces which act on the electrons which constitute it. The force exerted on a unit charge is $E + \frac{1}{c}[VH]$, where V is the velocity of the charge relative to the system of instruments used to measure the forces. The reason for this particular specification of V will be clearer later (see Chaps. XIII and XIV). It will be observed that it is not inconsistent with the deduction from the magnetic properties of electric currents which led us to include the term $\frac{1}{c}[VH]$ in the expression for the force on a charged body (p. 114). So far as any evidence which has been considered up to the present is concerned, we might as well have taken V to be the velocity of the charge relative to the aether, which we might suppose to be absolutely fixed in space. When we come to consider the electrical and optical properties of bodies in very rapid motion, we shall see that the assumption that V refers to the velocity relative to the measuring system effects very important simplifications.

The charge present in the element of volume $d\tau$ being $\rho\, d\tau$, the force exerted on this element of volume will be

$$dF = \rho\left(E + \frac{1}{c}[VH]\right)d\tau,$$

and the force on the whole volume will be given by the vector equation

$$F = \iiint \rho \left(E + \frac{1}{c} [VH] \right) d\tau$$

$$= \iiint \left\{ \operatorname{div} E \,.\, E + \frac{1}{c} [\rho V, H] \right\} d\tau,$$

since $\rho = \operatorname{div} E$. But

$$\operatorname{rot} H = \frac{1}{c} \left\{ \frac{\partial E}{\partial t} + \rho V \right\};$$

hence $$\rho V = c \operatorname{rot} H - \frac{\partial E}{\partial t},$$

and $$F = \iiint \left\{ \operatorname{div} E \,.\, E + \frac{1}{c} [c \operatorname{rot} H \,.\, H] - \frac{1}{c} \left[\frac{\partial E}{\partial t} \,.\, H \right] \right\} d\tau.$$

But $$\left[\frac{\partial E}{\partial t} \,.\, H \right] = \frac{\partial}{\partial t} [EH] - \left[E \,.\, \frac{\partial H}{\partial t} \right]$$

$$= \frac{\partial}{\partial t} [EH] + c [E \,.\, \operatorname{rot} E],$$

$$[E \,.\, \operatorname{rot} E] = -[\operatorname{rot} E \,.\, E],$$

and $$\operatorname{div} H = 0;$$

2. hence $$F = -\iiint \underline{\frac{1}{c} \frac{\partial}{\partial t}} [EH] \, d\tau + \iiint \{ \operatorname{div} E \,.\, E + [\operatorname{rot} E \,.\, E] \} \, d\tau$$

$$+ \iiint \{ \operatorname{div} H \,.\, H + [\operatorname{rot} H \,.\, H] \} \, d\tau \quad \text{...(2)}.$$

This is the total force on the volume τ. Consider the x component of F due to the third term of (2). Call it X_H. Then

$$X_H = \iiint \left\{ \left(\frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y} + \frac{\partial H_z}{\partial z} \right) H_x + H_z \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \right) - H_y \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) \right\} d\tau$$

$$= \iiint \left\{ H_x \frac{\partial H_x}{\partial x} - H_y \frac{\partial H_y}{\partial x} - H_z \frac{\partial H_z}{\partial x} + H_x \frac{\partial H_y}{\partial y} + H_y \frac{\partial H_x}{\partial y} + H_z \frac{\partial H_x}{\partial z} + H_x \frac{\partial H_z}{\partial z} \right\} d\tau$$

$$= \iiint \left\{ \frac{\partial}{\partial x} [\tfrac{1}{2} (H_x^2 - H_y^2 - H_z^2)] + \frac{\partial}{\partial y} (H_x H_y) + \frac{\partial}{\partial z} (H_x H_z) \right\} d\tau$$

$$= + \iint dS \{ l [\tfrac{1}{2} (H_x^2 - H_y^2 - H_z^2)] + m H_x H_y + n H_x H_z \},$$

where l, m, n are the direction cosines of the outward drawn normal to the element of area dS of the enclosing surface.

The third term of (2) will give rise to similar expressions for the components of force parallel to the y and z axes. If we treat the second term in (2) in the same way we shall get similar expressions except that the components of the magnetic intensity are replaced by the corresponding components of the electric intensity. If we put

$$p_{xx} = \tfrac{1}{2}\{E_x^2 - E_y^2 - E_z^2 + H_x^2 - H_y^2 - H_z^2\},\quad p_{yz} = p_{zy} = E_yE_z + H_yH_z,$$
$$p_{yy} = \tfrac{1}{2}\{E_y^2 - E_z^2 - E_x^2 + H_y^2 - H_z^2 - H_x^2\},\quad p_{zx} = p_{xz} = E_zE_x + H_zH_x,$$
$$p_{zz} = \tfrac{1}{2}\{E_z^2 - E_x^2 - E_y^2 + H_z^2 - H_x^2 - H_y^2\},\quad p_{xy} = p_{yx} = E_xE_y + H_xH_y,$$

the components of the last two terms of (2) may be written, selecting the x component as an example, in the form

J

$$\underline{\iint (lp_{xx} + mp_{yx} + np_{zx})\, dS.}$$

The last two terms therefore reduce to surface integrals over the boundary and are, in fact, identical with the forces due to the Maxwell stresses discussed in Chap. II, except for the addition of the magnetic terms which were not then being considered.

M

We now notice a very important difference between our present problem and the static case considered in Chap. II. The resultant of the forces acting on the volume from without, as calculated from the Maxwell stresses across the boundary, is no longer equal to the force tending to accelerate the charges enclosed by the boundary, the former being the greater by

$$\iiint \frac{1}{c}\frac{\partial}{\partial t}[EH]\, d\tau.$$

Thus, in the absence of electric charges, the resultant force due to the stresses over the boundary does not vanish unless the value of this volume integral is zero. The most natural interpretation would seem to be the following. In static cases, considering the action between the charges enclosed and the region external to the bounding surface, the resultant force on the external region, given by the Maxwell stresses, is equal and opposite to the resultant force on the charges enclosed. In general, however, the action and reaction between the charges and the external region are not equal and opposite; but part of the time considered and T' is an interval of time so great that it may be regarded as infinitely large.

It is evident that the electromagnetic momentum is relatively unimportant except when rapid changes are taking place in the state of the electromagnetic field.

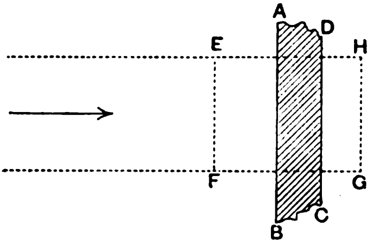


Fig. 29.

Let us now apply our results to the case of a plane polarized wave of monochromatic light incident normally on a surface which absorbs it completely. Let $ABCD$ represent the absorbing substance, the light being incident on it in the direction of the arrow which is parallel to the z axis. Let the electric intensity in the light wave lie along the axis of x and be equal to $E_0 \cos pt$. Then the components of the electric and magnetic intensities are

$$E_x = E = E_0 \cos pt, \quad E_y = 0, \; E_z = 0,$$

$$H_x = 0, \quad H_y = E_x = E_0 \cos pt = H = E, \quad H_z = 0.$$

Applying equation (2) to the cylinder whose cross section by the plane of the paper is $EFGH$ and integrating over a complete period, we see that the part coming from the volume integral on the right-hand side vanishes. The left-hand side is equal to the average force exerted by the light on the matter in $ABCD$ by which it is absorbed. The components of the Maxwell stresses are :—

$$p_{xx} = \tfrac{1}{2}\{E_x^2 - E_y^2 - E_z^2 + H_x^2 - H_y^2 - H_z^2\} = 0,$$

$$p_{yy} = \tfrac{1}{2}\{E_y^2 - E_z^2 - E_x^2 + H_y^2 - H_z^2 - H_x^2\} = 0,$$

$$p_{zz} = \tfrac{1}{2}\{E_z^2 - E_x^2 - E_y^2 + H_z^2 - H_x^2 - H_y^2\} = -E^2,$$

$$p_{xy} = p_{yx} = E_x E_y + H_x H_y = 0.$$

Similarly $$p_{yz} = p_{xz} = 0.$$

Thus there is no stress on the cylindrical surface of which EH and FG are sections, since this surface is everywhere parallel to the axis of z. The stress vanishes over GH since there is no light there. The only part of the surface over which the Maxwell stresses are effective is the end EF, and they are here equivalent to a pressure E^2 per unit area. Thus the force exerted by the light on the opaque surface is a pressure which is equal per unit area to

$$\frac{1}{T}\int E^2\,dt = \frac{1}{T}\int E_0^2 \cos^2 pt\,dt = \frac{E_0^2}{2}.$$

The energy per unit volume of the beam of light is

$$\tfrac{1}{2}\{E^2 + H^2\} = E^2,$$

so that the intensity of the light, which is equal to the average amount of energy transported across unit area in unit time, is

$$\frac{1}{T}\int_0^T cE^2\,dt = \frac{cE_0^2}{2}. \quad = P \cdot c$$

Thus in the case of light incident normally on a perfectly absorbing surface the radiation pressure is equal to the intensity of the light divided by its velocity of propagation.

It is clear that in the case of a perfectly reflecting surface the value of the radiation pressure will be doubled since the *intensity* of the light at EF will be twice as great as with a perfectly absorbing surface, when the incident intensity is the same. Since these results are independent of the plane of polarization of the light they will also be true, at normal incidence, when the light is unpolarized. It follows, on similar grounds, that they are also true for light of mixed frequencies.

Isotropic Radiation.

Next consider a perfectly reflecting enclosure filled with isotropic radiation. By isotropic radiation we mean radiation which is being propagated in all directions in such a manner that the probability of the direction of propagation of any ray, selected at random, being found within a given solid angle is proportional to that solid angle. The thermal radiation which would fill the enclosure in the final state of equilibrium which ensues when there are material bodies within it is of this character (see Chap. XV). Consider an infinitesimal area dS of the reflecting enclosure and apply equation (2) to a cylinder, similarly situated to that in Fig. 29, but whose height is infinitesimal compared with the dimensions of its ends. As before, let dS be perpendicular to the axis of z. We may now neglect the tractions on the sides of the cylinder on account of proportional to one another, differing only by the constant factor β. It will therefore be necessary for us to calculate only one of them.

Let us now denote our moving system by S and suppose that corresponding to S there is a fixed system S_1 which is obtained by enlarging the dimensions of the moving system S in the ratio $1 : \sqrt{1-\beta^2}$ in the direction of motion. Every point x, y, z in the moving system will correspond to a point x_1, y_1, z_1 in the fixed system, where the relation between x, y, z and x_1, y_1, z_1 is that already given. Also, corresponding elements of volume will contain equal charges, since straining the dimensions does not alter the quantity of electricity present. Thus, if ρ is the volume density in the moving system and ρ_1 the corresponding quantity in the corresponding element of volume in the fixed system, we have the equation

$$\rho\, dx\, dy\, dz = \rho_1\, dx_1\, dy_1\, dz_1.$$

But $dx = dx_1$, $dy = dy_1$, and $dz = dz_1\sqrt{1-\beta^2}$;

hence $$dx\, dy\, dz = dx_1\, dy_1\, dz_1\sqrt{1-\beta^2},$$

and $$\rho_1 = \rho\sqrt{1-\beta^2} \quad \text{.......................(3).}$$

The potential ϕ_1 of the distribution in the fixed system must satisfy Poisson's equation

$$\frac{\partial^2\phi_1}{\partial x_1^2} + \frac{\partial^2\phi_1}{\partial y_1^2} + \frac{\partial^2\phi_1}{\partial z_1^2} = -\rho_1 = -\sqrt{1-\beta^2}\,.\,\rho \quad \text{.........(4).}$$

But the scalar potential ϕ which we are seeking, satisfies the equation

$$\frac{\partial^2\phi}{\partial x_1^2} + \frac{\partial^2\phi}{\partial y_1^2} + \frac{\partial^2\phi}{\partial z_1^2} = -\rho,$$

whence $$\phi_1 = \sqrt{1-\beta^2}\,.\,\phi \quad \text{.......................(5).}$$

Thus the scalar potential in the moving system is equal to $1/\sqrt{1-\beta^2}$ times the electrostatic potential in the corresponding fixed system.

It is desirable to emphasize at this point that the fixed system which we have imagined is simply a mathematical device to facilitate the calculations. We are not supposing that the moving system is transformed physically in any way into the corresponding fixed system. What we have proved is that the scalar and

The results we have obtained are true for any electrical system in uniform motion. We shall apply them to two simple cases which we shall select partly on the ground of their simplicity but also because we may reasonably expect that they will give us an insight into the behaviour of the electron. The case that we shall consider first is that of a rigid thin spherical shell of uniform electrification.

The Rigid Electron.

We shall now suppose that the moving system consists of a spherical shell of radius R which is uniformly electrified. The thickness of the shell will be supposed to be negligible compared with its radius. It follows from the results that have already been established, that the energy, momentum, etc. of the field due to such an electrified shell when in motion may be obtained if we can calculate the static potential ϕ_1 of the corresponding system which arises when the actual space of the problem is strained so that all lengths parallel to the direction of motion are increased in the ratio $1 : \sqrt{1-\beta^2}$, whilst lengths at right angles to this remain unchanged. The moving sphere will evidently strain into an ellipsoid of revolution in the corresponding fixed system. The axis of revolution of the ellipsoid coincides with the direction of motion of the sphere. The major axis of any of the principal elliptic sections is equal to $R/\sqrt{1-\beta^2}$, whilst the minor axis is equal to the radius R of the moving sphere.

It is to be observed that these results will only be true provided the electrified sphere is rigid. Certain experiments, which will be discussed later, have led physicists to suspect that the lengths of bodies depend on their velocities relative to that of the observer engaged in measuring them. If this kind of change affects the electron itself, as well as the aggregate of electrons which we suppose constitutes the material substance, the figure in the fixed system which corresponds to the moving sphere will no longer be the ellipsoid which we have described. This follows because the moving system which we suppose to be spherical when at rest becomes distorted, and is no longer spherical when in motion. We shall see later that if the shape of the charged sphere does change, and if the changes are such as would naturally be suggested by the results of the experiments, the calculations charged body, when it moves with the velocity u, carries along with it an amount of electromagnetic momentum m_0u and of electrokinetic energy $\frac{1}{2}m_0u^2$. Moreover this momentum and energy remain unaltered so long as the velocity of the body is unaltered. But this behaviour is precisely what characterizes the motion of a so-called material particle whose mass is m_0. It is true that our results would have to be modified for particles moving with velocities comparable with the velocity of light. But they would nevertheless be exact, within the limits of accuracy of measurement, for such relative velocities as have been imparted to any considerable material masses in the universe. On the other hand the differences should be perceptible in the case of the very rapidly moving charged particles emitted by the radioactive substances, and, as we shall see, it is precisely the properties of these particles which have confirmed the results of the electromagnetic theory.

The idea of electromagnetic inertia, which is due to J. J. Thomson*, is fundamental to the electron theory of matter. For it opens up the possibility that the mass of all matter is nothing else than the electromagnetic mass of the electrons which certainly form part, and perhaps form the whole, of its structure. It obviously opens up the possibility of an electrical foundation for dynamics. This will be considered later.

Our calculations so far have presupposed that the moving charged body possesses and has always possessed a constant velocity in a straight line. A fuller discussion of electromagnetic mass involves the consideration of bodies undergoing acceleration and for such cases the results which we have obtained are not strictly true. This difficulty is one which is peculiar to the electromagnetic theory and arises from the fact that when a charged body is accelerated part of its energy travels off to infinity in the form of electromagnetic radiation.

Longitudinal and Transverse Mass.

We shall, however, see in the next chapter that, provided the acceleration of the body is sufficiently small compared with its velocity, the values for the energy in the field and for the electromagnetic momentum, which we have calculated above, will only differ from the true values by amounts which are exceedingly small. Under these circumstances, which will be stated more precisely later, the formulae which we have obtained for constant velocities may be applied to systems moving with varying velocities. The state of systems which satisfy this condition has been called by Abraham* quasi-stationary.

---

* *Phil. Mag.* vol. XI. p. 229 (1881). The idea that the mass of ordinary matter is of this character on account of the electrons [illegible] contains appears to have been first suggested by Larmor (*Phil. Trans.* vol. CLXX[illegible]. 697, 1895).

Assuming that the quasi-stationary condition is satisfied we shall now consider the behaviour of the moving electrical system under the influence of an accelerating force.

There are two cases to consider: (1) when the force is in the direction of motion and (2) when it is perpendicular to it. Any other case may be compounded from these two. We shall suppose the moving charge to be placed in an external field whose action gives rise to the accelerations under consideration. If we consider any infinitely distant surface enclosing the whole electrical system, the Maxwell stresses over it will vanish, so that the force exerted by the external field on the moving charge will be equal to the rate of *diminution* of the electromagnetic momentum of the external field. But the total momentum of the whole system remains constant, so that the external force must be equal to the rate of *increase* of the momentum of the moving charge. Since this is, by hypothesis, massless it follows that the force exerted on the moving charge by the external field is equal to the rate of increase of the electromagnetic momentum of the charge. This equality is clearly a vectorial one and is therefore true for the different components of the force and the momentum independently of one another.

In the case of a force acting in the direction of motion we evidently have

$$F_x = \frac{\partial}{\partial t}(G_x) = \frac{\partial G_x}{\partial \dot{x}} \frac{\partial \dot{x}}{\partial t},$$

since G_x only contains t implicity through $\dot{x}$. Comparing this with the equation $F_x = m\frac{\partial \dot{x}}{\partial t}$ we see that the mass for longitudinal accelerations is

$$\frac{\partial G_x}{\partial \dot{x}} = \frac{1}{c}\frac{\partial G}{\partial \beta} \quad \text{.......................(26).}$$

---

* *Loc. cit.*

When the force is transverse to the direction of motion the body will begin to describe a circular orbit with constant speed. If r is the instantaneous radius of curvature of the path, the rate of change of the momentum G is always directed along r and, by the principle of the hodograph, is equal per unit time to Gu/r, where u is the instantaneous velocity. Hence

$$F_y = \frac{G}{r} u.$$

But the instantaneous acceleration is also along r and equal to u^2/r, whence it follows that the electromagnetic mass for transverse accelerations is equal to

$$\frac{G}{r} u \div \frac{u^2}{r} = \frac{G}{u} = \frac{1}{c} \frac{G}{\beta} \quad \text{..................(27).}$$

The mass for transverse accelerations is therefore different from the mass for longitudinal accelerations. This difference was first pointed out by Abraham*. It happens, as may easily be verified, that the difference between the longitudinal and the transverse electromagnetic mass becomes vanishingly small for small velocities.

The properties of an electrically charged body which we have been considering have a close analogy in hydrodynamics. Any geometrical figure moving in a fluid sets the surrounding fluid in motion. In the steady state when the figure moves uniformly in a straight line the fluid motion is carried along by the moving figure as though it were rigidly attached to it; when the state of motion changes, waves are set up and part of the energy of the system is radiated away to great distances. We shall see in the next chapter that this also has its counterpart in the electrical case. Confining ourselves to the case of uniform motion, in the steady state it is found† that if the moving figure is intrinsically massless it nevertheless possesses inertia and behaves as though it had a certain mass coefficient which is a function of the mass of the fluid displaced by it. In the case of a massless sphere moving in a perfect fluid this apparent mass is one-half of that of the fluid displaced by the sphere. When a circular cylinder moves at right angles to its length the apparent mass is equal to that of the fluid displaced by the cylinder. In the case of unsymmetrical figures the apparent mass is no longer the same for motions in different directions. The hydrodynamical cases differ from the electrical ones in one important respect; the apparent mass is always independent of the velocity of the moving figure.

---

* *Loc. cit.*

† Lamb's *Hydrodynamics*, pp. 85, 130.

This analogy was very clearly seen by J. J. Thomson in the original memoir* in which he developed the idea of electromagnetic mass. We shall take the liberty of quoting his exact words: "The charged sphere will produce an electric displacement throughout the field; and as the sphere moves, the magnitude of this displacement at any point will vary. Now, according to Maxwell's theory, a variation in the electric displacement produces the same effect as an electric current; and a field in which electric currents exist is a seat of energy; hence the motion of the charged sphere has developed energy, and consequently the charged sphere must experience a resistance as it moves through the dielectric. But as the theory of the variation of the electric displacement does not take into account anything corresponding to resistance in conductors, there can be no dissipation of energy through the medium; hence the resistance cannot be analogous to an ordinary frictional resistance, but must correspond to the resistance theoretically experienced by a solid in moving through a perfect fluid. In other words, it must be equivalent to an increase in the mass of the charged moving sphere."

The Contractile Electron.

We have seen that the determination of the field due to a rigid spherical shell of electrification in motion can be reduced to the determination of the electrostatic potential due to a certain ellipsoid. We have pointed out already that the particular ellipsoid which we have been led to consider as the equivalent fixed system depends upon our supposition that the spherical shell is rigid. Now the negative results which have been obtained in a number of optical experiments on moving systems, instituted largely in order to try to detect relative motion between the system and the luminiferous medium, seem incapable of explanation except on the hypothesis, suggested by FitzGerald†, that, on account of the motion, the matter of the testing system undergoes the influence of the electrostatic field alone, a group of particles all having the same velocity would be deflected to the right or to the left according to the direction of the field. The spot would be shifted from the centre of the photographic plate. Since the deflection depends upon the velocity, if the group consisted of particles having different velocities the spot would be spread out into a line lying in the plane of the paper. Each point on this line would correspond to a particular velocity.

---

* *Phil. Mag.* V. vol. II. p. 230 (1881).

† *Nature*, June 16 (1892).

The lines of force of the magnetic field run from left to right in the plane of the paper, so that if it alone were operative it would cause the particles which pass through O and D to describe circles in a plane perpendicular to the plane of the paper. Owing to the magnetic field, then, the spot would be displaced in the plane perpendicular to the plane of the paper, if all the particles had the same velocity, and would be drawn out into a line in this plane, if the particles had different velocities. When both fields are operative at once we should expect to get a curved line on the photographic plate, each point of which represents the point of impact of particles having a certain velocity. The position of each point gives us, of course, the magnetic and electrostatic deflection of a particle with a definite but unknown velocity. It remains to be seen how we may deduce from the measured displacements the value of the mass as a function of the velocity of the moving particle.

For the sake of simplicity we shall suppose a uniform electric field to extend, from left to right, all the way from O to D and then to cease absolutely. The magnetic field H is uniform and parallel to this electric field all the way from O to the photographic plate P.

Taking O as the origin, let the coordinates of D be 0, 0, z, and those of points on the plate P be x, y, z'. Let the axis of y lie in the plane of the paper, Y denoting the electric intensity. Assuming that the deflections may be treated as small, we have, so far as the motion in the plane of the paper is concerned,

$$m\frac{\partial^2 z}{\partial t^2} = 0 \text{ and } m\frac{\partial^2 y}{\partial t^2} = Ye, \text{ from } O \text{ to } D,$$

and

$$m\frac{\partial^2 y}{\partial t^2} = 0, \text{ from } D \text{ to } P.$$

Thus $$\frac{\partial z}{\partial t} = \text{const.} = w_0,$$

where w_0 is the component along OD of the velocity of projection,

and $$z = w_0t + \text{const.} = w_0t$$

if $t = 0$ when the particle leaves O.

Hence $t = z/w_0$ and

$$Ye = m\frac{\partial^2 y}{\partial t^2} = mw_0^2\frac{\partial^2 y}{\partial z^2}.$$

Thus $$mw_0^2 y = \tfrac{1}{2}Yez^2 + Az + B.$$

Since $y = 0$ when $z = 0$ and also when $z = z_1$,

$$B = 0 \text{ and } A = -\tfrac{1}{2}Yez_1.$$

Thus $$mw_0^2 y = \tfrac{1}{2}Yez(z - z_1),$$

$$\frac{\partial y}{\partial z} = \frac{1}{mw_0^2}(Yez - \tfrac{1}{2}Yez_1),$$

and $$\left(\frac{\partial y}{\partial z}\right)_{z=z_1} = \frac{1}{mw_0^2}(\tfrac{1}{2}Yez_1) = \tan\theta,$$

where θ is the angle which the tangent to the parabolic path at D makes with OD. Since the subsequent path is rectilinear the displacement y' at the photographic plate is

$$y' = (z' - z_1)\tan\theta = \frac{1}{2mw_0^2}Yez_1(z' - z_1).$$

The projection of the path on the plane of xOz is controlled by the magnetic field and is a circle passing through O and D to the degree of approximation of this calculation. Let r be the radius of this circle and $S_1D_1P_1$ (Fig. 30 *b*) the tangent to it at the point S_1 which is symmetrical with respect to O and D. D_1 and P_1 are the intersections with horizontal planes through D and P respectively. $P'(x', 0, z')$ is the intersection of the projected trajectory on the photographic plate. S' is the intersection of a vertical line $P'S'$ and a horizontal line S_1SS'. S is the mid-point of OD. Then

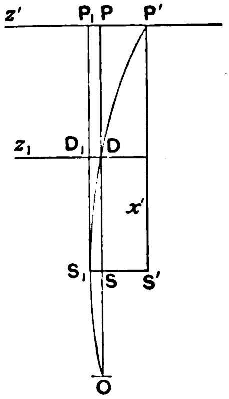


Fig. 30 *b*.

$$P_1P = S_1S = \frac{OS^2}{2r - SS_1} = \frac{z_1^2}{8r},$$

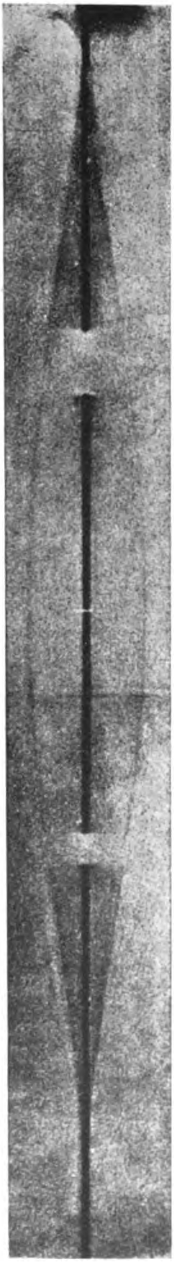


Fig. 32.

the plane of the paper. A photographic film extends all the way round the inside of the cylinder. The whole apparatus is exhausted and is placed in a uniform magnetic field H perpendicular to the plane of the paper.

To simplify the discussion of the experiment we shall suppose the distance between A and B to be so small that it is quite negligible. In the absence of the electric and magnetic fields the β-rays travel in straight lines all the way from R and their trace on the photographic film is a circle in the plane of the condenser. This circle becomes a straight line, of course, when the film is unrolled. When the fields are applied the paths become more complex. Between the plates they are still horizontal straight lines, but after escaping they describe circular spirals with their axes along the direction of the magnetic force. The rays will only be able to escape from the plates provided the downward pull of the magnetic force is equal to the upward pull of the electric field. If we consider a particle starting out with velocity w in a direction making an angle α with that of the magnetic field, the condition for compensation is

$$Xe = Hew \sin \alpha,$$

where X is the electric intensity between the plates. Thus for the particles to escape from the condenser

$$\beta = \frac{w}{c} = \frac{X}{Hc \sin \alpha}.$$

In Bucherer's experiments this formula was tested by taking X and H so that $X/Hc = \frac{1}{2}$. Under these circumstances

$$\sin \alpha = 1/2\beta.$$

After leaving the plates the particles follow a spiral path of which ST may be regarded as

At the outset we are face to face with a new difficulty which arises from the fact that the charge $\rho d\tau$ which enters into the integrals expressing the potentials is not all present in the element of volume at the same instant. Thus the charge $\rho d\tau$ which occurs in this element of the integral is not equal to the true charge which would occur in this element if it were at rest. Consider the truncated cone *TOSR* whose apex is at *P* and which is terminated by the spherical surfaces *AOSC* and *BTRD* whose radii are $c\theta$ and $c(\theta + d\theta)$ respectively. The sphere *AC* represents the instantaneous position of contributing charges at the instant $t - \theta$ and the sphere *BD* the instantaneous position at $t - \theta - d\theta$.

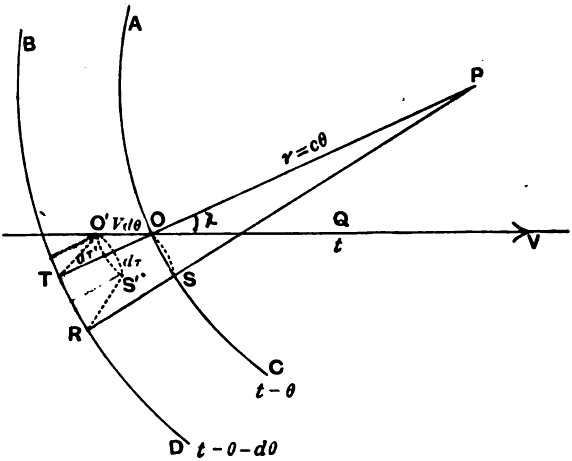


Fig. 33.

Thus the part of the charge which contributes to the potential at *P* from the front-end *OS* of the element of volume $d\tau$ is present there at a later time than that which contributes from the back-end *TR*. All the charges are, however, present in the displaced element $d\tau' = O'S'RT$ at the one instant $t - \theta - d\theta$ provided $OO' = Vd\theta$. Hence the true charge de which is effective in the element $d\tau$ is, if λ is the angle between r and V,

$$de = \rho d\tau' = \rho d\tau (1 - \beta \cos \lambda)^*.$$

---

* This result is due to Wiechert, *Arch. Néerl.*, (2) vol. v. p. 549 (1900). Cf. also A. Liénard, *L'Éclairage électrique*, vol. xvi. pp. 5, 53, 106 (1898).

We have

$$\phi = \frac{e}{4\pi}\frac{1}{c\theta - \frac{1}{c}\{(x-\xi)\dot{\xi} + (y-\eta)\dot{\eta} + (z-\zeta)\dot{\zeta}\}},$$

$$-\frac{\partial\phi}{\partial x}$$

$$= \frac{ec}{4\pi}\frac{c^2\frac{\partial\theta}{\partial x} - \dot{\xi} + \dot{\xi}\frac{\partial\xi}{\partial x} + \dot{\eta}\frac{\partial\eta}{\partial x} + \dot{\zeta}\frac{\partial\zeta}{\partial x} - (x-\xi)\frac{\partial\dot{\xi}}{\partial x} - (y-\eta)\frac{\partial\dot{\eta}}{\partial x} - (z-\zeta)\frac{\partial\dot{\zeta}}{\partial x}}{[c^2\theta - \{(x-\xi)\dot{\xi} + (y-\eta)\dot{\eta} + (z-\zeta)\dot{\zeta}\}]^2}.$$

But $$\frac{\partial\xi}{\partial x} = \frac{d\xi}{d(t-\theta)}\frac{\partial(t-\theta)}{\partial\theta}\frac{\partial\theta}{\partial x} = -\dot{\xi}\frac{\partial\theta}{\partial x}.$$

Similarly

$$\frac{\partial\eta}{\partial x} = -\dot{\eta}\frac{\partial\theta}{\partial x}, \quad \frac{\partial\zeta}{\partial x} = -\dot{\zeta}\frac{\partial\theta}{\partial x}, \quad \frac{\partial\dot{\xi}}{\partial x} = -\ddot{\xi}\frac{\partial\theta}{\partial x},$$

$$\frac{\partial\dot{\eta}}{\partial x} = -\ddot{\eta}\frac{\partial\theta}{\partial x} \text{ and } \frac{\partial\dot{\zeta}}{\partial x} = -\ddot{\zeta}\frac{\partial\theta}{\partial x}.$$

Hence

$$-\frac{\partial\phi}{\partial x}$$

$$= \frac{ec}{4\pi}\frac{-\dot{\xi} + \{c^2 - (\dot{\xi}^2 + \dot{\eta}^2 + \dot{\zeta}^2) + (x-\xi)\ddot{\xi} + (y-\eta)\ddot{\eta} + (z-\zeta)\ddot{\zeta}\}\frac{\partial\theta}{\partial x}}{[c^2\theta - \{(x-\xi)\dot{\xi} + (y-\eta)\dot{\eta} + (z-\zeta)\dot{\zeta}\}]^2},$$

and

$$c\frac{\partial\theta}{\partial x} = \frac{1}{r}\left\{x - \xi - (x-\xi)\frac{\partial\xi}{\partial x} - (y-\eta)\frac{\partial\eta}{\partial x} - (z-\zeta)\frac{\partial\zeta}{\partial x}\right\},$$

or $$\frac{\partial\theta}{\partial x} = \frac{x-\xi}{c^2\theta - \{(x-\xi)\dot{\xi} + (y-\eta)\dot{\eta} + (z-\zeta)\dot{\zeta}\}}.$$

Thus

$$-\frac{\partial\phi}{\partial x} = \frac{ec}{4\pi}\frac{\begin{array}{l}[c^2\{(x-\xi) - \theta\dot{\xi}\} + \dot{\xi}\{(x-\xi)\dot{\xi} + (y-\eta)\dot{\eta} + (z-\zeta)\dot{\zeta}\} \\ \quad - (x-\xi)\{\dot{\xi}^2 + \dot{\eta}^2 + \dot{\zeta}^2\} \\ \qquad + (x-\xi)\{(x-\xi)\ddot{\xi} + (y-\eta)\ddot{\eta} + (z-\zeta)\ddot{\zeta}\}]\end{array}}{(c^2\theta - \{(x-\xi)\dot{\xi} + (y-\eta)\dot{\eta} + (z-\zeta)\dot{\zeta}\})^3}.$$

The expressions for $-\frac{\partial\phi}{\partial y}$ and $-\frac{\partial\phi}{\partial z}$ may be written down from inspection, on interchanging the axes.

The expressions for E_y and E_z may be written down by interchanging the axes.

It will be observed that of the three terms in the expression for E_x the first does not involve the acceleration $\ddot{\xi}, \ddot{\eta}, \ddot{\zeta}$. Let us consider this part of the electric intensity separately. Its components are proportional to

$$(x - \xi - \theta\dot{\xi}), \quad (y - \eta - \theta\dot{\eta}) \text{ and } (z - \zeta - \theta\dot{\zeta})$$

respectively. The resultant of this part of the electric intensity is therefore proportional to and directed along O_1P, where O_1 is the position which the moving charge would occupy if it continued from the instant $t - \theta$ to move uniformly during the following interval θ with the velocity $\dot{\xi}, \dot{\eta}, \dot{\zeta}$ which it had at that instant. We may therefore write the resultant E_1 of the part of the electric intensity which does not depend upon the acceleration in the form

$$E_1 = \frac{e(1 - \beta^2)}{4\pi r^3 (1 - \beta\cos\lambda)^3} O_1P \quad \text{.................(3)}.$$

If we work out the value of the magnetic intensity from the expression $H = \text{rot } U$ we find that there is a part of that also which is independent of the acceleration. Denoting this by H_1 we find

$$H_1 = \beta E_1 \sin\lambda_1 \quad \text{.......................(4)},$$

where λ_1 is the angle between O_1P and the direction of V. H_1 is tangential to the circle passing through P in the plane perpendicular to V and whose centre is on the direction of V.

In the case of a particle which moves with a uniform velocity, the expressions just given will represent the whole of the electric and magnetic intensities. In that case it is convenient to express $E_1 = E$ in terms of the instantaneous radius r_1 and the angle it makes with the direction of motion, rather than in terms of r and λ. The change is easily made. We have

$$r_1/\sin\lambda = r/\sin\lambda_1$$

and since

$$OO_1 = V\theta = \beta r,$$

$$r_1^2 = r^2 + \beta^2 r^2 - 2\beta r^2 \cos\lambda$$

$$= r^2(1 - \beta\cos\lambda)^2 + \beta^2 r^2 \sin^2\lambda$$

$$= r^2(1 - \beta\cos\lambda)^2 + \beta^2 r_1^2 \sin^2\lambda_1;$$

that they have the point O_0 as centre. When this is done the volume of the shell between any two spheres is the same as when they were excentric. The total energy in the field is clearly the same as if the energy of each excentric shell was spread uniformly over the volume of the corresponding concentric shell. It is therefore

$$\frac{e^2}{24\pi}\frac{3+\beta^2}{1-\beta^2}\int_R^\infty \frac{cd\theta}{r^2} = \frac{e^2}{24\pi}\frac{3+\beta^2}{1-\beta^2}\int_R^\infty \frac{dr}{r^2}$$

$$= \frac{e^2}{24\pi R}\frac{3+\beta^2}{1-\beta^2} = \frac{e^2}{8\pi R}\left\{1+\frac{4\beta^2}{3(1-\beta^2)}\right\} \quad \ldots(8).$$

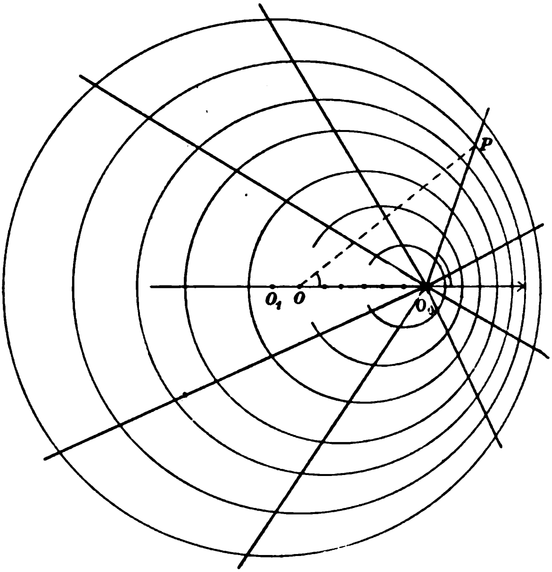


Fig. 35.

In our units the potential (electrostatic) energy outside the sphere of radius R is $e^2/8\pi R$. The remainder must therefore be the energy of the magnetic field of the moving charge. Since the electrostatic energy is the same whether the particle is at rest or in motion, the magnetic energy is the same thing as the kinetic energy of the moving charge. The kinetic energy is thus

$$\frac{e^2}{6\pi R}\frac{\beta^2}{1-\beta^2} \quad \ldots\ldots\ldots\ldots(9).$$

the components of E_2, and P_2, Q_2, R_2 may readily be obtained since we know that E_2 and H_2 are equal to each other and are also perpendicular both to each other and to OP. The value of this part of the energy per unit volume is

$$[X_1X_2 + Y_1Y_2 + Z_1Z_2 + P_1P_2 + Q_1Q_2 + R_1R_2].$$

Introducing the element of volume in polar coordinates we integrate this throughout the region between the two shells and find

$$dU_2 = \frac{1}{(1-\beta^2)^2} \frac{e^2}{3\pi r} \frac{\beta\Gamma}{c} \cos\epsilon \, d\theta \quad \ldots\ldots\ldots\ldots(13).$$

This energy is a maximum when $\epsilon = 0$ and is zero when $\epsilon = \pi/2$. Since $\Gamma d\theta$ is the change of velocity in the interval $d\theta$ we see that $\Gamma d\theta \cos\epsilon$ is equal to the resolved part of the change of velocity in the direction of motion. Calling this $c\delta\beta$ we have

$$dU_2 = \frac{e^2}{3\pi r(1-\beta^2)^2} \beta d\beta \ldots\ldots\ldots\ldots\ldots\ldots(14).$$

This part of the energy contains r in the denominator, so that it falls to zero as the two spheres proceed to infinity.

There is a rather simple and important relation between the energy dU_2 and the energy in the field when the charge is in uniform motion. In considering this, in order to fix our ideas, we shall suppose that the motion is uniform except during the infinitesimal interval between $t-\theta$ and $t-\theta+d\theta$. Before $t-\theta$ the charge moves in a straight line with uniform velocity v_1 and after $t-\theta+d\theta$ it again moves uniformly in a straight line but with a different velocity v_2. Consider the energy in the field outside of the two excentric spheres. Up to the instant t this is that which is proper to the case of a charge moving with the uniform velocity v_1. At the instant t this begins to be altered. A new field begins to be established, namely that which is proper to a particle moving with the uniform velocity v_2. The new field is evidently left behind it as the excentric shell travels outwards with the velocity c. The relation referred to is that the energy dU_2 is just what is required to change the field *outside* of the excentric sphere from the state corresponding to the uniform velocity v_1 to the state corresponding to v_2. For this reason the energy dU_2 has, very appropriately, been called by Langevin the wave of reorganization.

the field due to a point charge e, which lay outside a sphere of radius r about the charge. It is very unlikely that an electron possesses this constitution, and where the electrification has an extended geometrical configuration the equation has to be modified. We have seen that the second term on the right-hand which arises from the energy radiated is not affected by the geometrical arrangement of the charge. With the first term it is different. The new value of this may be found by considering the difference in the energy in the steady fields corresponding to the velocities before and after the action of the force. Proceeding in this way we find, in the case of any distribution possessing spherical symmetry, that r is replaced by the radius of the sphere multiplied by a numerical factor. The precise value of the factor depends on the configuration of the distribution, whether superficial or throughout the volume of the sphere and, if so, whether uniform or not. Thus the form of the equation is unaltered even when the electric charge is not concentrated at a point.

Our equation differs from the corresponding Newtonian Equation $F = m\Gamma$ in two important respects. In the first place we have the additional term depending upon Γ^2 and in the second place m is never constant, although it is approximately so when β is very small. This aspect of the question has already been considered. When the acceleration is small the second term containing Γ^2 will be relatively unimportant, so that the Newtonian law is an approximation which is true for small accelerations and small velocities.

For the *form* of the Newtonian law to be preserved it is necessary that the second term should be negligible compared with the first. Thus if a is the radius of the electron it is necessary that

$$\frac{\delta U_3}{\delta U_2} = \frac{(1 - \beta^2 \sin^2 \epsilon)\,\Gamma a}{2\,(1 - \beta^2)\,c^2 \beta \cos \epsilon} \qquad \text{...............(16)}$$

should be small compared with unity. Since $1 - \beta^2$ may in all practical cases be taken to be of the order unity, the order of magnitude of this fraction is

$$\frac{\Gamma a}{2c^2 \beta \cos \epsilon} = \frac{1}{2c}\,\Gamma\,\frac{a}{c\beta \cos \epsilon} = \frac{1}{2c}\,\Gamma\,\frac{a}{v \cos \epsilon} = \frac{\Gamma \tau}{2c} \text{......(17)},$$

where τ is the time necessary for the charged sphere to move in the direction of the acceleration through a distance equal to its own radius. It is therefore necessary that the change in the velocity of the body during this time τ should be small compared with the velocity of radiation.

If this condition is satisfied the motion will be the same as that calculated on the supposition that radiation does not occur. Abraham[*] has suggested that the term *quasi-stationary* be applied to states of motion which satisfy this condition. Since the calculation of the motion when the state is not quasi-stationary is extremely complicated, it is important to inquire whether such states are likely to occur in nature or not. There is one case which is of very frequent occurrence in which the condition is violated, and that is the case in which the velocity of the particle and β are zero. This state, however, only lasts for an insignificant interval of time in the case of an electron. We know from the value of the electromagnetic mass of these particles that a is about 10^{-13} cm. If the particle has moved from rest for a time t the velocity $v = \Gamma t$, and since $c = 3 \times 10^{10}$ cms. per sec. the fraction $\dfrac{\Gamma a}{2cv \cos \epsilon}$ will be comparable with $10^{-13}/(6 \times 10^{10} \times t)$. This will clearly be negligible compared with unity for any measurable interval of time. Even after 10^{-20} sec. it will have fallen almost to 10^{-4}. Thus even in the case of a charged particle starting from rest it seems unlikely that serious error will arise if the motion is treated as though it were quasi-stationary.

Another case in which one might expect the quasi-stationary condition to be departed from arises when the acceleration is very great. Probably the greatest acceleration with which we are familiar is that which occurs during the impact of a β-ray particle on an atom. There is some evidence that in favourable cases a β-ray particle may be completely stopped within a distance d equal to the diameter of an atom, let us say 10^{-8} cm. Assuming uniform acceleration as an illustration, the equation of energy in this case is

$$\tfrac{1}{2} m v^2 = m \Gamma d.$$

Thus
$$\Gamma = v^2/(2 \times 10^{-8}),$$

and
$$\frac{\Gamma a^{-}}{2cv} = \tfrac{1}{4}\,(10^{-5})\,\frac{v}{c} = \tfrac{1}{4} \times 10^{-5},$$

---

[*] *Ann. der Phys. loc. cit.*

when v has its maximum possible value c. Even in this case the ratio has a value which is quite small, so that it does not seem likely that the assumption of quasi-stationary motion will lead to serious error in practical cases.

The Reaction of the Radiation.

Although the part of the force which is neglected by the assumption of quasi-stationary motion is practically always small compared with the remaining part, there are cases in which it might exert important effects through its persistency. For example in the case of periodic motions, if the reaction of the radiation always tended to stop the motion, it would in time exert an appreciable effect. For this reason it is desirable to calculate the reaction on an accelerated electron without assuming, as we did at the beginning of this chapter, that the increase in the velocity during the time required by the electron to move over its own radius is small compared with the velocity of radiation. Owing to the difficulties which arise in a more general treatment we shall content ourselves with the case in which the velocity, though variable, is rectilinear, and in which the squares and higher powers of the ratios of the velocity of the electron, and the derivatives of the velocity, to the velocity of light may be neglected.

Let us seek the force acting on an element of volume $d\tau'$ of the moving charged sphere at the point P whose coordinates are x', y', z' at the instant t. Let the variable velocity u of the sphere be parallel to the axis of x. Let $Q\ (x, y, z)$ be the position of some other point of the sphere at the time t. The part of the potentials at P which arises from the element $d\tau = dx\,dy\,dz$ which is at the point Q at the time t will be determined, not by the instantaneous state and position Q of the element of volume, but by its state and position $Q_\theta = x_\theta, y, z$ at a time $t - \theta$, where

$$PQ_\theta = \{(x_\theta - x')^2 + (y - y')^2 + (z - z')^2\}^{\frac{1}{2}} = c\theta.$$

We also have

$$PQ = \{(x - x')^2 + (y - y')^2 + (z - z')^2\}^{\frac{1}{2}} = r,$$

and

$$x_\theta = x - \frac{\partial x}{\partial t}\theta + \frac{\partial^2 x}{\partial t^2}\frac{\theta^2}{2} - \frac{\partial^3 x}{\partial t^3}\frac{\theta^3}{6} + \dots$$

$$= x - u\theta + \tfrac{1}{2}\dot{u}\theta^2 - \tfrac{1}{6}\ddot{u}\theta^3 + \dots.$$

Since $c\theta$ and r differ only by terms involving u and its derivatives as a factor we can put $r = c\theta$ in the equation for x_θ.

Thus $$x_\theta = x - \frac{u}{c} r + \frac{\dot{u}}{2c^2} r^2 - \frac{\ddot{u}}{6c^3} r^3 + \dots .$$

Substituting this value of x_θ in PQ_θ, we obtain

$$PQ_\theta^2 = c^2\theta^2 = (y - y')^2 + (z - z')^2 + \left\{\frac{u}{c} r - \frac{\dot{u}}{2c^2} r^2 + \frac{\ddot{u}}{6c^3} r^3 - \dots - (x - x')\right\}^2$$

$$= r^2 - 2(x - x')\left(\frac{u}{c} r - \frac{\dot{u}}{2c^2} r^2 + \frac{\ddot{u}}{6c^3} r^3 - \dots\right)$$

+ terms involving squares of small quantities.

Thus $$c\theta = r\left\{1 - 2\frac{x - x'}{r}\left(\frac{u}{c} - \frac{\dot{u}}{2c^2} r + \frac{\ddot{u}}{6c^3} r^2 - \dots\right)\right\}^{\frac{1}{2}}$$

$$= r\left\{1 - \frac{x - x'}{r}\left(\frac{u}{c} - \frac{\dot{u}}{2c^2} r + \frac{\ddot{u}}{6c^3} r^2 - \dots\right)\right\}$$

+ higher terms.

We also have

$$\frac{\partial x_\theta}{\partial x} = 1 - \frac{u}{c}\frac{\partial r}{\partial x} + \frac{\dot{u}}{c^2} r \frac{\partial r}{\partial x} - \frac{\ddot{u}}{2c^3} r^2 \frac{\partial r}{\partial x} + \dots .$$

But $$\frac{\partial r}{\partial x} = \frac{x - x'}{r},$$

so that

$$dx_\theta = \left(1 - \frac{u}{c}\frac{x - x'}{r} + \frac{\dot{u}}{c^2}(x - x') - \frac{\ddot{u}}{2c^3} r (x - x') + \dots\right) dx.$$

Thus to every element of volume $d\tau \, (= dx\,dy\,dz)$ of the charge at the time t there will be the corresponding element

$$d\tau_\theta \, (= dx_\theta \, dy \, dz)$$

at the time θ seconds previously, where

$$d\tau_\theta = \left(1 - \frac{u}{c}\frac{x - x'}{r} + \frac{\dot{u}}{c^2}(x - x') - \frac{\ddot{u}}{2c^3} r (x - x') + \dots\right) d\tau.$$

The velocity of this charge at the time $t - \theta$ is

$$u - \dot{u}\theta + \tfrac{1}{2}\ddot{u}\theta^2 - \tfrac{1}{6}\ddot{u}\theta^3 + \dots$$

$$= u - \dot{u}\frac{r}{c} + \tfrac{1}{2}\ddot{u}\frac{r^2}{c^2} - \tfrac{1}{6}\ddot{u}\frac{r^3}{c^3} + \dots .$$

In integrating for the scalar and vector potentials we must replace $\frac{d\tau}{r}$ by $\frac{d\tau_\theta}{c\theta}$. On substituting the values already obtained, we find to the first order of small quantities that

$$\frac{d\tau_\theta}{c\theta} = \frac{d\tau}{r}\left\{1 + \frac{\dot{u}}{2c^2}(x - x') - \frac{\ddot{u}}{3c^3}(x - x')\,r + \dots\right\}.$$

The scalar potential

$$\phi = \frac{1}{4\pi}\iiint\left(\rho\,\frac{d\tau}{r}\right)'$$

is evidently

$$= \frac{1}{4\pi}\iiint\frac{\rho}{r}\left\{1 + \frac{\dot{u}}{2c^2}(x - x') - \frac{\ddot{u}}{3c^3}(x - x')\,r + \dots\right\}d\tau.$$

The vector potential

$$U_x = \frac{1}{4\pi c}\iiint\left(\rho\,\frac{u\,d\tau}{r}\right)'$$

contains u as a factor, so that the replacement of $\frac{d\tau}{r}$ by $\frac{d\tau_\theta}{c\theta}$ will only introduce small quantities of the second order and may be omitted. It is, however, necessary to substitute in U_x the value of u at the instant $t - \theta$. Thus

$$U_x = \frac{1}{4\pi c}\iiint\frac{\rho}{r}\left\{u - \frac{\dot{u}}{c}\,r + \tfrac{1}{2}\,\frac{\ddot{u}}{c^2}\,r^2 - \dots\right\}d\tau.$$

$U_y = U_z = 0$, since the resultant velocity is parallel to the axis of x.

The x component of the force on a unit charge at P is

$$F_x = -\frac{\partial\phi}{\partial x'} - \frac{1}{c}\,\frac{\partial}{\partial t}(U_x) + \frac{1}{c}\,[uH].$$

The magnetic force H (= rot U) is derived by differentiating the vector potential with respect to the coordinates. It thus contains u as a factor, so that the lowest term in $[uH]$ will involve squares of small quantities. This term may therefore be neglected. Evidently

$$\frac{\partial\phi}{\partial x'} = -\frac{1}{4\pi}\iiint\frac{\rho}{r^2}\,\frac{\partial r}{\partial x'}\,d\tau + \frac{\dot{u}}{8\pi c^2}\iiint\rho\left\{-\frac{x - x'}{r^2}\,\frac{\partial r}{\partial x'} - \frac{1}{r}\right\}d\tau$$
$$+ \frac{\ddot{u}}{12\pi c^3}\iiint\rho\,d\tau + \dots$$
$$= \frac{1}{4\pi}\iiint\rho\,\frac{x - x'}{r^3}\,d\tau + \frac{\dot{u}}{8\pi c^2}\iiint\rho\left\{\frac{(x - x')^2}{r^3} - \frac{1}{r}\right\}d\tau$$
$$+ \frac{\ddot{u}}{12\pi c^3}\iiint\rho\,d\tau + \dots,$$

This is
$$F_x' u\,dt = -\frac{e^2}{6\pi a c^2} u\dot{u}\,dt + \frac{e^2}{6\pi c^3} u\ddot{u}\,dt$$
$$= -\frac{e^2}{6\pi a c^2}\frac{d}{dt}(\tfrac{1}{2}u^2)\,dt + \frac{e^2}{6\pi c^3}\left\{\frac{d}{dt}(u\dot{u}) - \dot{u}^2\right\}dt.$$

The term $-\frac{e^2}{6\pi a c^2}\frac{d}{dt}(\frac{1}{2}u^2)\,dt$ represents the diminution of the electromagnetic energy ($\frac{1}{2}mv^2$) of the steady part of the field of the moving charged body. The term $-\frac{e^2}{6\pi c^3}\dot{u}^2 dt$ is equal to the energy radiated.

The work done by the second term in F_x' in a finite time, say from t_1 to t_2, is
$$\frac{e^2}{6\pi c^3}\left\{\Big|u\dot{u}\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2}\dot{u}^2 dt\right\}.$$

The integrated part vanishes if either u or $\dot{u}$ is zero at the limits. In the case of periodic motions this term will vanish when the limits are the recurrent zero values of u and $\dot{u}$. Even if these limits are not chosen its value over a long period of time will be small, being comparable with the maximum value within a single period. The value of $\dot{u}^2$, however, is always positive, when not zero, so that the integral $\int_{t_1}^{t_2}\dot{u}^2 dt$ will increase indefinitely as t_2 increases. It will evidently tend continuously to diminish the kinetic energy of the vibrating electrified particle, and thus to stop its motion. It is, in fact, very similar in effect to the action of a frictional force on a dynamical system.

Planck* has suggested that this force may account for the frictional term which it is necessary to introduce in order to account for absorption in optics. Whilst there is no doubt that this force must be operative it is, I think, far too small to account for any of the observed effects. The kind of action in which the frictional term is smallest is found in the case of substances which give rise to the residual rays. We have seen in Chapter VIII that even in these cases the coefficient which enters into the resistance term is of the order 10^{12} in the units employed, whereas the term $\frac{e^2\ddot{u}}{6\pi c^3}$ would lead to only about 10^{-7} of this amount. The main part of what appears to be a damping force in optical radiators must evidently be sought elsewhere.

---

* *Vorlesungen über die Theorie der Wärmestrahlung*, p. 109. Leipzig (1906).

the material which is necessary to determine the specific effects which are caused by the motion of material media.

It appears to the writer to be impossible to acquire a true perspective of the matter at issue without considering it in its historical development. We shall therefore take it up from that point of view, at the risk perhaps of lengthening the treatment. Naturally, optical effects will be treated as a particular case of electrodynamic actions.

The Aberration of Light.

The study of this subject arose out of a discovery by Bradley in 1728, made during an investigation whose object was to detect annual parallax in certain fixed stars near the zenith. Such a parallax was found. It was not, however, directed towards the sun as it would have been if it were ordinary stellar parallax, but it was in a direction perpendicular to this in the plane of the earth's orbit. The magnitude of the "aberration" was found to be proportional to the sine of the inclination of the star, but was constant for stars of equal inclination. The results were found to be capable of complete explanation on the view that the light was propagated in space with a finite velocity in a direction which was fixed relatively to the star and which was uninfluenced by the earth's motion.

The problem is one of relative motion and can be made quite clear by considering an analogous material case. Suppose an observer to be in an open carriage which is moving with uniform velocity in a horizontal plane. He wishes to determine the direction of motion of the drops which fall into the carriage in a rain storm. To do this he is provided with a long tube which can rotate about a horizontal axis perpendicular to the length of the tube, and the inclination can be read off on a circle in a vertical plane. The direction is then determined by adjusting the tube so that the drops fall through without reaching the sides. The direction of the rain as thus determined will evidently depend on the direction of the rain relative to axes fixed in the ground, on its velocity and on that of the carriage relative to the same axes. Let us consider the comparatively simple case when the

It should therefore vary directly as the index of refraction of the medium. An experiment to test this conclusion was carried out long afterwards by Airy who used a telescope filled with water. He was, however, surprised to find that under these circumstances the constant of aberration had the same value as in an ordinary telescope.

An experiment based on somewhat similar reasoning had occurred independently to Arago, who argued that since the deviation of a ray of light produced by a dense prism depends on the ratio of the velocity of light in the prism and in the surrounding medium, its magnitude ought to be different according to whether the passage of the light through the prism is helped or retarded by the motion of the earth. With this experiment it is evidently unnecessary to use light from extra-terrestrial sources. Arago found that no effect due to the earth's motion could be detected, although the expected effect was comparatively large.

An explanation of the experimental results of both Airy and Arago was given in a general way by Fresnel, who suggested that the aether was carried along by moving material media in such a way as to compensate exactly for the difference between the velocities of the light in the medium and in vacuo. It is necessary that the aether should be carried along with a velocity which is only a fraction of that of the medium, for if it were carried with the same velocity there would be no aberration at all. Fresnel's suggestion was worked out more completely, later, by Maxwell and Stokes. We shall now calculate with what fraction of the velocity of the transparent medium it is necessary that the aether should be carried along in order to give the result required by Arago's experiment that refraction is independent of the earth's motion.

The path of a ray of light in any medium is determined by the fact that the time required to pass from any one point of the path to any other has a stationary value. This extension of Fermat's Principle of Least Time follows on the undulatory theory of light from the fact that if A and B are any two points in a ray the disturbance arising from points in the wave front in the immediate neighbourhood of A must all reach B in the same phase. This condition will evidently be satisfied when the time from A to B has either a stationary or a maximum value as well as when the value is a minimum. Thus if ds is an element of the path of the ray and V' the velocity at any point, then

$$\delta\int\frac{ds}{V'}=0 \quad\text{..............................(1).}$$

Consider the case of any optical system, including the observer, which moves through space with the velocity whose components are u, v, w. Let l, m, n be the direction cosines of the ray at any point. Suppose that in any refracting medium the velocity of the light passing through it is increased by the amount θ multiplied by the velocity of this medium through the aether. If V is the standard value of the velocity of light in this medium when the system is at rest, the velocity relative to axes fixed in space for the moving system will be

$$V+\theta(lu+mv+nw).$$

But the observer is moving relative to the fixed axes with a velocity whose components are u, v, w, so that the resolved part of his velocity along the direction of the ray is $lu+mv+nw$ and the velocity of the ray relative to him is

$$V-(1-\theta)(lu+mv+nw).$$

The equation of the relative ray paths is therefore

$$\delta\int\frac{ds}{V-(1-\theta)(lu+mv+nw)}=0,$$

or to the first order of $(u, v, w)/V$

$$\delta\int\frac{ds}{V}+\delta\int\frac{1-\theta}{V^2}(udx+vdy+wdz)=0.$$

To be in agreement with Arago's experimental result it is necessary that the relative paths to this order should be independent of u, v, w. This will be the case if the quantity under the second integral sign is a perfect differential; since the value of the integral will then depend only on the terminal points which are not varied. If m is the refractive index of the medium and c is the velocity of light in free aether $m^2=c^2/V^2$. It is therefore necessary and sufficient that

$$\frac{m^2(1-\theta)}{c^2}(udx+vdy+wdz)$$

be a perfect differential. Since the relation between m and u, v, w

a change in the velocity of light in the region around the periphery of a massive iron flywheel when it was made to rotate rapidly. The results were entirely negative. A way of escape from most of the objections to Stokes's theory has recently been suggested by H. A. Wilson* (see p. 283).

The Michelson and Morley Experiment.

So far those who held the view of an aether which is undisturbed by the motion of matter through it appeared to have much the better of the argument. The foregoing account represents the state of the question when Michelson and Morley carried out the famous experiment by which they hoped to obtain positive evidence of the relative motion of the earth and the aether. Before describing this experiment we shall first consider the problem of the reflection of light at a moving surface.

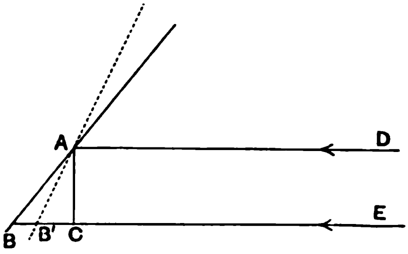


Fig. 38.

The plane mirror AB is moving towards the right with the uniform velocity v. A plane wave of light bounded by the rays DA, EB falls on it. Let us find the relation between the angle of incidence ι and the angle of reflection. When a wave meets the mirror at A on the ray DA, it has reached the point C on the ray EB. If the mirror were at rest this wave would afterwards meet it at B, but as the mirror is moving, the wave will meet it at B', where $BB'/B'C = v/c$, c being the velocity of light. The mirror thus behaves as though it were turned through an angle ϵ given by

$$\frac{\tan(\iota-\epsilon)}{\tan\iota}=\frac{B'C}{BC}=\frac{c}{c+v}=\frac{1}{\tan\iota}\frac{\tan\iota-\tan\epsilon}{1+\tan\iota\tan\epsilon},$$

---

* *Phil. Mag.* VI. vol. xix. p. 809 (1910).

The corresponding difference of path is

$$c(T+T'-T_1)=l\frac{v^2}{c^2}.$$

If the fringes were adjusted when the apparatus was stationary a displacement corresponding to this path difference should be observed if the whole apparatus were moving with uniform velocity v in the direction indicated. If the motion were with uniform velocity in a direction perpendicular to this the path difference is found to be equal but in the opposite direction; thus if the apparatus is adjusted so that *SD* coincides with the earth's motion relative to the sun, the effect of turning it through a right angle should be to cause the fringes to shift by an amount corresponding to a path difference equal to $2l\frac{v^2}{c^2}$.

Michelson and Morley set up their apparatus on a stone slab which floated in mercury. It could thus be rotated without causing strains to be set up in the apparatus. As the rotation was made to take place the fringes were observed continuously but no displacement of their position could be detected. The investigators were able to show conclusively that the relative motion of their apparatus and the aether could not amount to as much as one sixth part of the velocity of the earth in its orbit.

The result of the Michelson and Morley experiment was to place the problem of the relative motion of matter and aether in an apparently impossible position. On the one hand the view maintained by Stokes that the aether was carried along by the earth in its motion appeared to be full of inconsistencies, on the other hand the consequences of the stagnant aether hypothesis were flatly contradicted by experiment. An escape from the dilemma was pointed out by Fitzgerald[*] who suggested that the null effect in the Michelson and Morley experiment was due to a change in the dimensions of the apparatus in the direction of the earth's motion, just sufficient to counterbalance the expected effect. The required change would evidently be produced if the matter of the apparatus contracted in the direction of the earth's motion in the ratio $1-v^2/2c^2$ to 1, the lengths in planes perpendicular to this direction being unaltered. This change would be too small to be capable of direct measurement in any case; but even if it were not, the operation could not be carried out; as any material scale which might be used would also contract, in the same ratio as the material to be measured.

This hypothesis seems a wild speculation at first sight, but it appears, on further inquiry, that it is rather what might be expected to occur if matter is made up of electrons. In that case the question at once suggests itself as to why a given portion of solid matter preserves its shape. If the matter were made up of superposable elements of positive and negative electrification, capable of infinitesimal subdivision, the only state of stable equilibrium would be one in which any excess of charge of one sign would be dissipated to infinity and the remaining equal and opposite charges would be superposed on each other. The matter would thus annihilate itself and disappear. In order that matter should be stable enough to preserve its identity it is necessary that the ultimate elements of electrification should be finite, and it is also necessary that the superposition of opposite elements should not result in annihilation. To ensure this it is necessary that the ultimate elements of opposite sign should not have exactly equal geometrical distributions. We conclude, in fact, that the positive and negative electrons are essentially different. The simplest assumption we can make as to the nature of the forces which keep them in equilibrium is to suppose that they are under the influence of equal and opposite forces of ultimately electrical character, but it may be that this supposition is inadequate and that the electrical forces are balanced by forces of non-electrical type. In any event, in static cases the equilibrium configuration of the electrons will be determined by the positions in the matter at which the resultant force is zero.

We know from the results of Chapter XI that the field due to an electric system in motion differs from that due to the same system at rest, in such a way as would result if all the lengths parallel to the direction of motion were changed in the ratio of $(1-v^2/c^2)^{\frac{1}{2}}$ to 1. Thus for the electrical forces to have the same value in the moving as in the fixed system it is necessary that all lengths in the former which are parallel to the direction of motion should be reduced in the ratio of $(1 - v^2/c^2)^{\frac{1}{2}}$ to 1. If the forces are all electrical it follows that the positions of the balancing points where the force vanishes will all be nearer together in this ratio and a contraction in length of this amount, in the direction of motion, will have to occur, in the moving matter, if equilibrium is to be maintained. Since

$$(1 - v^2/c^2)^{\frac{1}{2}} = 1 - v^2/2c^2$$

as far as the fourth order in v/c, we see that this contraction is of the right magnitude to account for the absence of an effect in the Michelson and Morley experiment. Another mode of explanation will be developed later.

An attempt has been made by Trouton and Rankine (*R. S. Proc.* A, vol. LXXX. p. 420, 1908) to detect and measure the Fitzgerald shrinkage by measuring the resistance of a metal strip when (*a*) parallel, and (*b*) perpendicular, to the earth's motion. The experiment showed that if any shrinkage occurs it is compensated to within 2 per cent. of the expected value by some counterbalancing effect of the motion on the resistance. Trouton and Rankine were able to show that on the electron theory of metallic conduction the changes in the mass, mean free path and velocity of the electrons would exactly compensate the effect of the expected shrinkage. It is interesting to note that according to their calculation the Lorentz change of electromagnetic mass with velocity gives the desired compensation, whereas the value found by Abraham does not.

Besides those which have already been discussed, there are a number of other cases, where effects due to motion through the aether might be expected to arise, which have engaged the attention of physicists. As is well known, the phenomenon of ordinary double refraction can be fully accounted for solely by the fact that the media which exhibit it transmit the disturbance which constitutes light with different velocities in different directions. We have seen that in order to explain Arago's experimental result it is necessary to suppose that the velocity of light relative to the observer, through a refracting medium which is isotropic when at rest, depends on the motion of the latter through the aether. The relative velocity of the wave will therefore be different in different directions, and we might expect double refraction to occur when the medium is moving. Such an effect has been very carefully sought for but always with negative results.

Another effect which has been looked for is a change in the rotation of the plane of polarization of plane polarized light produced by its passage through a quartz plate, due to change in the relative direction of the apparatus and the earth's motion. The rotation of the plane of polarization of light produced by its passage through quartz in a direction near that of the optic axis is believed to be due to the difference of velocity of circularly polarized rays when travelling in that direction. Thus the quartz is able to increase the velocity of one of the two oppositely directed circularly polarized rays, into which the plane polarized ray may be resolved, relatively to the other. The emergent ray, produced by their recombination, is thus polarized in a plane different from that of the incident ray. Since the velocity of either of the circularly polarized rays in the moving quartz depends on a function of its refractive index for that ray multiplied by the velocity of the quartz through the aether, and the refractive index is different for the two rays, an effect should arise which is proportional to v/c. Experiment shows that there is no effect which is comparable with that to be expected on this view. The reader will find a number of other experiments and observations, which bear upon the subject under discussion, described in the last chapter of Whittaker's *History of the Aether and Electricity*.

A method of reconciling Stokes's theory of the optical properties of moving bodies with the experimental facts, without assuming the Fitzgerald contraction, has recently been indicated by H. A. Wilson*. Wilson points out that the problem of the motion of a body like the earth, through the aether, may be solved, in such a way that the motion of the incompressible aether is everywhere continuous and irrotational, provided that the tangential relative velocity vanishes at the surface of the body. All the conditions may be satisfied by an appropriate flow of the aether

---

* *Phil. Mag.* VI. vol. xix. p. 809 (1910).

Referring to Chapter VIII, p. 148, we see that the polarization P is related to the electromotive intensity E' by the equation

$$P = (m^2 - 1) E'.$$

When the medium is in motion, with velocity V,

$$E' = E + \frac{1}{c} [V.H],$$

if the magnetic permeability of the medium is unity, in accordance with equation (5), Chapter IX. In the present case, therefore,

$$P_x = (m^2 - 1) \left(E_x - \frac{w}{c} H_y \right) \quad \ldots\ldots\ldots\ldots\ldots(2),$$

if we take the axis of x parallel to the electric intensity and that of y parallel to the magnetic intensity and the motion of the medium parallel to the axis of z with velocity w. The current density is $\dot{D}_x + w \frac{\partial P_x}{\partial z}$, the latter term arising from the convection of the polarization by the moving medium. Thus the fundamental electromagnetic equations become

$$-\frac{\partial E_x}{\partial z} = \frac{1}{c} \dot{H}_y \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3),$$

and $$-\frac{\partial H_y}{\partial z} = \frac{1}{c} \left(\dot{D}_x + w \frac{\partial P_x}{\partial z} \right) = \frac{1}{c} \left(\dot{E}_x + \dot{P}_x + w \frac{\partial P_x}{\partial z} \right) \quad \ldots(4).$$

Eliminating P and H from (4) by means of (2) and (3) we have

$$\left(1 - (m - 1) \frac{w^2}{c^2} \right) \frac{\partial^2 E_x}{\partial z^2} = \frac{m^2}{c^2} \frac{\partial^2 E_x}{\partial t^2} + 2 \frac{w}{c^2} (m^2 - 1) \frac{\partial^2 E_x}{\partial t \partial z}.$$

If V is the velocity of light in the moving dielectric with respect to the fixed aether we shall have $E_x = Ae^{ip(t - z/V)}$, where A and p are constants. Substituting this value and neglecting w^2/c^2 we have

$$c^2 = m^2 V^2 - 2w (m^2 - 1) V,$$

or $$V = \frac{c}{m} + \left(1 - \frac{1}{m^2} \right) w \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(5).$$

This is Fresnel's formula (p. 273). This deduction shows that the convection of the polarized dielectric through the fixed aether produces the same change in the velocity of the light as is required by the experiments of Airy and Fizeau and by Fresnel's hypothesis. Those experiments therefore do not prove that the aether is precisely the same function of

$$x'' = x - ut,\ y' = y,\ z' = z,\ t'' = t - \beta^2 \frac{u}{c} x''$$

that E_x is of x, y, z, t. Thus in comparing a moving and a fixed system, which are identical with one another when at relative rest, for any event in the fixed system characterised by a certain set of values of E and H, there is in the moving system a corresponding event characterised by values of E' and H', which are the same functions of x'', y', z' and t'' that E and H are of x, y, z and t. It follows that within a self-contained system, to the order of u/c, the electromagnetic effects are independent of the velocity of the system, since E' and H' are the values of the electric and magnetic forces which would be actually measured by instruments moving along with the system. This is only true, however, provided the time recorded by clocks in the moving system is the local time and not the "true" time recorded by clocks in the fixed system.

The principle of correlation is due to Lorentz. In the first instance he only succeeded in establishing it to the first order of u/c after the manner of the discussion above. A little later Larmor by including a contraction along the axes of motion, to accord with Fitzgerald's hypothesis, showed that the principle held to the order u^2/c^2. In 1903 Lorentz showed that a very similar transformation enabled the correlative principle to be established with exactness for all values of the velocity u less than that of light. We shall now investigate what is the necessary transformation of the electromagnetic equations in order to establish correspondence up to any order of u/c.

Since the differential equations are linear the transformation we are seeking will be a linear one and the preceding investigation at once suggests the form which is most likely to be successful. Let x, y, z, t be the coordinates referred to the first set of axes and x', y', z', t' those referred to the second, which are in motion relatively to the first. Assume that both x' and t' are linear functions of both x and t whilst y' and z' involve only y and z respectively. Let us put

$$x' = i(x + jt),\ y' = ky,\ z' = lz,\ t' = m(t + nx) \ldots\ldots(10),$$

$$\left.\begin{aligned}-m\left(1-\frac{j^2}{c^2}\right)\frac{\partial H_x}{\partial t'} &= ck\frac{\partial}{\partial y'}\left(E_z-\frac{j}{c}H_y\right)-ck\frac{\partial}{\partial z'}\left(E_y+\frac{j}{c}H_z\right)\\ -m\frac{\partial}{\partial t'}\left(H_y-\frac{j}{c}E_z\right) &= ck\frac{\partial E_x}{\partial z'}-ci\frac{\partial}{\partial x'}\left(E_z-\frac{j}{c}H_y\right)\\ -m\frac{\partial}{\partial t'}\left(H_z+\frac{j}{c}E_y\right) &= ci\frac{\partial}{\partial x'}\left(E_y+\frac{j}{c}H_z\right)-ck\frac{\partial E_x}{\partial y'}\end{aligned}\right\}. \quad \ldots\ldots(15).$$

Now put $$j=-u,\ k=1,$$

and $$i=m=1\Big/\left\{m\left(1-\frac{j^2}{c^2}\right)\right\}=1\Big/\left(1-\frac{j^2}{c^2}\right)^{\frac{1}{2}}=1\Big/\left(1-\frac{u^2}{c^2}\right)^{\frac{1}{2}}=\beta.$$

Then

$$\left.\begin{aligned}\frac{1}{c}\left(\frac{\partial E_x}{\partial t'}+(u_1-u)\beta\rho\right) &= \frac{\partial}{\partial y'}\left\{\beta\left(H_z-\frac{u}{c}E_y\right)\right\}-\frac{\partial}{\partial z'}\left\{\beta\left(H_y+\frac{u}{c}E_z\right)\right\}\\ \frac{1}{c}\left(\frac{\partial}{\partial t'}\left\{\beta\left(E_y-\frac{u}{c}H_z\right)\right\}+v_1\rho\right) &= \frac{\partial H_x}{\partial z'}-\frac{\partial}{\partial x'}\left\{\beta\left(H_z-\frac{u}{c}E_y\right)\right\}\\ \frac{1}{c}\left(\frac{\partial}{\partial t'}\left\{\beta\left(E_z+\frac{u}{c}H_y\right)\right\}+w_1\rho\right) &= \frac{\partial}{\partial x'}\left\{\beta\left(H_y+\frac{u}{c}E_z\right)\right\}-\frac{\partial H_x}{\partial y'}\end{aligned}\right\} \quad \ldots\ldots(16),$$

$$\left.\begin{aligned}-\frac{1}{c}\frac{\partial H_x}{\partial t'} &= \frac{\partial}{\partial y'}\left\{\beta\left(E_z+\frac{u}{c}H_y\right)\right\}-\frac{\partial}{\partial z'}\left\{\beta\left(E_y-\frac{u}{c}H_z\right)\right\}\\ -\frac{1}{c}\frac{\partial}{\partial t'}\left\{\beta\left(H_y+\frac{u}{c}E_z\right)\right\} &= \frac{\partial E_x}{\partial z'}-\frac{\partial}{\partial x'}\left\{\beta\left(E_z+\frac{u}{c}H_y\right)\right\}\\ -\frac{1}{c}\frac{\partial}{\partial t'}\left\{\beta\left(H_z-\frac{u}{c}E_y\right)\right\} &= \frac{\partial}{\partial x'}\left\{\beta\left(E_y-\frac{u}{c}H_z\right)\right\}-\frac{\partial E_x}{\partial y'}\end{aligned}\right\} \ldots(17).$$

Thus if we put

$$E_x'=E_x,\quad E_y'=\beta\left(E_y-\frac{u}{c}H_z\right),\quad E_z'=\beta\left(E_z+\frac{u}{c}H_y\right),$$

$$H_x'=H_x,\quad H_y'=\beta\left(H_y+\frac{u}{c}E_z\right),\quad H_z'=\beta\left(H_z-\frac{u}{c}E_y\right),$$

$$u'=u_1-u,\quad v'=v_1/\beta,\quad w'=w_1/\beta,\quad \rho'=\beta\rho,$$

the equations for the new dependent variables in terms of the independent variables $x'=\beta(x-ut)$, $y'=y$, $z'=z$ and $t'=\beta(t-ux/c^2)$ are accurately identical with the differential equations (III) and (IV) connecting the corresponding undashed variables. It follows that the correlation previously established to the first order is true to any order provided the electromagnetic quantities have the undeformable axes of reference and an equally rigid unit of length. We may then fix the time of the occurrence by having a series of similar clocks distributed at infinitesimal distances apart in space and arranged so that when brought together they register equal times. So far we have not indicated any method by which the clocks may be correlated when they are not in the same place. We shall now assume that this can be done in such a way that the velocity of light in a vacuum measured by means of the system of axes and the clocks is equal to a universal constant c, which is independent of the state of rest or of uniform motion of the system. Thus if A and B are two material points whose distance apart is r, a ray of light emitted from A at the instant t_A will reach B at the instant t_B, where

$$r = c\,(t_B - t_A);$$

t_A and t_B are the readings of the clocks at A and B respectively. We shall assume, what is not of course *a priori* obvious, that this is true whatever the state of motion of the axes of reference may be, provided it is not accelerated. The assumption here involved has been called by Einstein the Principle of the Constancy of the Velocity of Light. It is a particular case of the Principle of Relativity which may be stated in the words: The laws of nature are independent of the state of motion of the system of reference provided this is unaccelerated.

Now consider two equivalent systems of axes S and S'. By equivalent we mean here such as possess equal units of length and clocks running at equal rates, when the two systems are at rest relatively to each other. Suppose that S' is in uniform motion relatively to S. It follows from the fundamental assumption that the velocity of light must have the same value when measured with reference to both S and S'. Consider any event occurring at the point x, y, z at the time t when referred to the system S. Referred to the system S' this event will be described by the corresponding variables x', y', z' and t'. It is required to determine the relations between x, y, z and t and x', y', z' and t'.

On account of the homogeneity of time and space we should expect these relations to be linear. It follows that the coordinate planes of S' will be in uniform motion when referred to those of S. In general they will not be perpendicular to each other. Let

Moreover, since the relation between y and y' or z and z' cannot depend on the sign of v,

$$\phi(v) = \phi(-v),$$

and therefore $\phi(v) = 1$, $\phi(v) = -1$ being obviously untrue.

We therefore obtain finally

$$\left.\begin{aligned} t' &= \beta\left(t - \frac{v}{c^2}x\right) \\ x' &= \beta(x - vt) \\ y' &= y, \quad z' = z \end{aligned}\right\} \quad \dots\dots\dots\dots(4).$$

It will at once be observed that these values are the same as those which we obtained in the last chapter by the direct transformation of the electromagnetic equations.

If these equations are solved for t in terms of t', etc., we obtain

$$\left.\begin{aligned} t &= \beta\left(t' + \frac{v}{c^2}x'\right) \\ x &= \beta(x' + vt') \\ y &= y', \quad z = z' \end{aligned}\right\} \quad \dots\dots\dots\dots(5),$$

showing that the system S is moving with respect to the system S' with velocity $-v$ along the axis of x'.

It follows from these equations that when a body originally at rest is set in motion, its dimensions, measured relative to axes in the original position of rest, are contracted in the direction of motion and unchanged in planes perpendicular to this. Let x_1', y_1', z_1' and x_2', y_2', z_2' be the coordinates of any two points in the body referred to axes moving with it (system S'), x_1, y_1, z_1 and x_2, y_2, z_2 being the coordinates of the same point referred to the system S relative to which the body is moving parallel to the x axis with velocity v. Then at any time t which is constant with reference to the system S we have

$$\left.\begin{aligned} x_1 - x_2 &= (1 - v^2/c^2)^{\frac{1}{2}}(x_1' - x_2') \\ y_1 - y_2 &= y_1' - y_2' \\ z_1 - z_2 &= z_1' - z_2' \end{aligned}\right\} \quad \dots\dots\dots\dots(6).$$

Thus a length equal to l' parallel to the axis of x' is reduced in the ratio $\sqrt{1 - \frac{v^2}{c^2}} : 1$ when measured with reference to axes with regard to which the body is moving with velocity v at a constant time t referred to these axes.

If instead of t constant we take t' the time referred to the axes moving with the body as constant, then

$$x_1' - x_2' = (1 - v^2/c^2)^{\frac{1}{2}} (x_1 - x_2) \text{.................} (7).$$

It follows that a body at rest with respect to S as measured from S' undergoes the same change of dimensions as a body at rest with respect to S' when measured from S. This result is required of course by the symmetry of the motion.

A similar set of relations holds with respect to the time. Suppose we have a clock moving with uniform velocity v along the axis of x when referred to the system S. Let us take the position of this clock as the origin of coordinates for the system S'. Then $x' = 0$ always and $x = vt$. Let t_1', t_2' be the times of two consecutive events as recorded on the moving clock and t_1, t_2 the times for the same events as registered in the fixed system S. Then

$$t_1' = \beta\left(t_1 - \frac{vx_1}{c^2}\right) = \beta t_1\left(1 - \frac{v^2}{c^2}\right),$$

so that $$t_1 = \beta t_1' \text{ and } t_2 = \beta t_2'.$$

Hence $$t_2 - t_1 = \beta (t_2' - t_1') \text{ ....} (8).$$

We may take t_2', t_1' to represent two consecutive strokes of the clock. It is then clear from the equation above that the moving clock as observed from the fixed system will appear to have its periodic time increased in the ratio $1 : \sqrt{1 - \frac{v^2}{c^2}}$. The frequency will be decreased in the inverse ratio.

This case may be realised physically by considering the line spectrum emitted by a moving molecule or ion. Measured with reference to axes at rest with respect to the ion the frequency of the emitted light in the case of many lines is confined within very narrow limits. The period of the light may thus be taken to represent that of the reference clock and t_2', t_1', etc. may be taken as the times at which the emitted vibrations are consecutively in the same phase. The above result shows that in addition to the well-known Doppler effect the frequency ν of the light given out by an ion or molecule moving with velocity v relative to the axes of each of which is less than that of light is also less than the velocity of light. For let $v = c - \lambda$ and $v_1' = c - \mu$, where λ and μ are positive and smaller than c, then the resultant velocity

$$v_1 = c\frac{2c - \lambda - \mu}{2c - \lambda - \mu + \lambda\mu/c},$$

and this is always less than c. Also the addition of any velocity to the velocity of light gives rise to a velocity which is still equal to the velocity of light.

Another interesting consequence of the foregoing results is that it is impossible for any signal to be transmitted from one point A to another B with a velocity W greater than c the velocity of light. Let the point A be taken at the origin of coordinates and the point B lie on the x axis at a distance l from A. Observers transmitting and receiving the signal are fixed at A and B respectively. Let the signal be transmitted by means of a material strip relatively to which it travels with velocity W in the direction $A \rightarrow B$. Now let the material strip carrying the signal be itself moving along the x axis with the velocity $-v$, where $v < c$, the velocity of light. The velocity of the signal relative to the transmitting and the observing system is evidently

$$\frac{W - v}{1 - Wv/c^2}.$$

The time required for the signal to be transmitted is thus

$$T = l\frac{1 - Wv/c^2}{W - v}.$$

Since v can have any value $< c$, T can always be made negative provided $W > c$. This would imply that the effect would be perceived before the cause had commenced to act. Although this may not necessarily involve any logical contradiction it is opposed to the whole character of our experience. The truth of the theorem therefore follows.

Some of the preceding results differ so considerably from those which follow from the generally accepted notions of space and time that many readers will probably regard them as serious objections to the views here developed. If, however, the principle of relativity is accepted they appear to follow with logical certainty. The *a priori* argument in favour of the principle of relativity will be considered later.

since
$$\sin^2\phi + \frac{\left(\cos\phi - \frac{v}{c}\right)^2}{1-\frac{v^2}{c^2}} = \left[\frac{\left(1-\frac{v}{c}\cos\phi\right)}{\sqrt{1-\frac{v^2}{c^2}}}\right]^2.$$

As the direction of the z axis is arbitrary this formula must be true generally.

Mechanics of an Electron or Material Point.

Let an electric charge e, of infinitesimal dimensions, move under the influence of an electromagnetic field. We shall assume, in accordance with the principle of relativity, that its equations of motion, referred to the system of axes S' with reference to which it is instantaneously at rest, are

$$\mu\frac{d^2x_0'}{dt_0'^2} = eE_x', \qquad \mu\frac{d^2y_0'}{dt_0'^2} = eE_y', \qquad \mu\frac{d^2z_0'}{dt_0'^2} = eE_z',$$

where μ *is a constant* which we shall call the mass of the electron. The suffix 0 is introduced to indicate that the moving point is instantaneously at rest with reference to the S' axes.

Let us now deduce the equations of motion of the point charge when referred to the system S relative to which S' is moving with velocity v. We have

$$dt_0' = \beta\left(dt_0 - \frac{v}{c^2}dx_0\right) = \beta\left(1-\frac{v\dot{x}_0}{c^2}\right)dt_0,$$

$$dx_0' = \beta\,(dx_0 - v\,dt_0) \quad = \beta\,(\dot{x}_0 - v)\,dt_0,$$

$$dy_0' = dy_0, \qquad dz_0' = dz_0.$$

Hence
$$\frac{dx_0'}{dt_0'} = \frac{\dot{x}_0 - v}{1-\frac{v\dot{x}_0}{c^2}},$$

$$\frac{d^2x_0'}{dt_0'^2} = \frac{d}{dt_0'}\left(\frac{\dot{x}_0 - v}{1-\frac{v\dot{x}_0}{c^2}}\right) = \frac{1}{\beta\left(1-\frac{v\dot{x}_0}{c^2}\right)}\frac{d}{dt_0}\left(\frac{\dot{x}_0 - v}{1-\frac{v\dot{x}_0}{c^2}}\right)$$

$$= \frac{1}{\beta}\,\frac{\left(1-\frac{v\dot{x}_0}{c^2}\right)\ddot{x}_0 + \frac{v}{c^2}(\dot{x}_0 - v)\,\ddot{x}_0}{\left(1-\frac{v\dot{x}_0}{c^2}\right)^3}. = \frac{1}{\beta^3}\,\frac{\ddot{x}_0}{\left(1-\frac{v\dot{x}_0}{c^2}\right)^3}$$

Similarly
$$\frac{dy_0'}{dt_0'} = \frac{1}{\beta\left(1-\frac{v\dot{x}_0}{c^2}\right)}\frac{dy_0}{dt_0},$$

$$\frac{d^2y_0'}{dt_0'^2} = \frac{1}{\beta\left(1-\frac{v\dot{x}_0}{c^2}\right)}\frac{1}{\beta}\frac{d}{dt_0}\left(\frac{\dot{y}_0}{1-\frac{v\dot{x}_0}{c^2}}\right)$$

$$= \frac{\left(1-\frac{v\dot{x}_0}{c^2}\right)\ddot{y}_0 + \dot{y}_0\frac{v}{c^2}\ddot{x}_0}{\beta^2\left(1-\frac{v\dot{x}_0}{c^2}\right)^3},$$

with a similar equation for $\frac{d^2z_0'}{dt_0'^2}$.

If S' is instantaneously at rest with reference to the moving point $\dot{x}_0 = v$, $\dot{y}_0 = 0$, $\dot{z}_0 = 0$, so that

$$\frac{d^2x_0'}{dt_0'^2} = \sqrt{1-\frac{v^2}{c^2}}\frac{\ddot{x}_0}{(1-v^2/c^2)^2} = \beta^3\ddot{x}_0,$$

$$\frac{d^2y_0'}{dt_0'^2} = \frac{\ddot{y}_0}{\beta^2(1-v^2/c^2)^2} = \beta^2\ddot{y}_0,$$

$$\frac{d^2z_0'}{dt_0'^2} = \frac{\ddot{z}_0}{\beta^2(1-v^2/c^2)^2} = \beta^2\ddot{z}_0.$$

Referring to equation (c), p. 307, we therefore get

$$\mu\beta^3\ddot{x}_0 = eE_x,$$

$$\mu\beta\ddot{y}_0 = e\left(E_y - \frac{v}{c}H_z\right),$$

$$\mu\beta\ddot{z}_0 = e\left(E_z + \frac{v}{c}H_y\right).$$

These are the equations which hold for the instant when $\dot{x}_0 = v$, $\dot{y}_0 = 0$, $\dot{z}_0 = 0$. We may on the left-hand side replace v by $q = \sqrt{\dot{x}_0^2 + \dot{y}_0^2 + \dot{z}_0^2}$ and on the right-hand by $\dot{x}_0$. Leaving out the suffix 0 and adding the other terms $\frac{\dot{y}}{c}H_z$, $-\frac{\dot{z}}{c}H_y$, the above equations may be written as a particular case of the symmetrical equations

$$\left.\begin{aligned} \frac{d}{dt}\left\{\frac{\mu\dot{x}}{\sqrt{1-q^2/c^2}}\right\} &= K_x \\ \frac{d}{dt}\left\{\frac{\mu\dot{y}}{\sqrt{1-q^2/c^2}}\right\} &= K_y \\ \frac{d}{dt}\left\{\frac{\mu\dot{z}}{\sqrt{1-q^2/c^2}}\right\} &= K_z \end{aligned}\right\} \quad \text{..................(16)},$$

where
$$\left.\begin{aligned}K_x &= e\left[E_x+\frac{\dot{y}}{c}H_z-\frac{\dot{z}}{c}H_y\right]\\ K_y &= e\left[E_y+\frac{\dot{z}}{c}H_x-\frac{\dot{x}}{c}H_z\right]\\ K_z &= e\left[E_z+\frac{\dot{x}}{c}H_y-\frac{\dot{y}}{c}H_x\right]\end{aligned}\right\}\ldots\ldots\ldots\ldots\ldots(17).$$

μ is a constant coefficient in all three equations and since

$$\frac{d}{dt}\left\{\frac{\mu\dot{x}}{\sqrt{1-q^2/c^2}}\right\}=\frac{\mu}{(1-q^2/c^2)^{\frac{3}{2}}}\left\{\ddot{x}\left(1-\frac{\dot{x}^2+\dot{y}^2+\dot{z}^2}{c^2}\right)+\dot{x}\,\frac{\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z}}{c^2}\right\},$$

with similar expressions for the terms in $\dot{y}$ and $\dot{z}$, it will be seen at once that equations (16) reduce to the equations on p. 310 for the particular case when $\dot{y}=\dot{z}=0$. Since the equations (16) and (17) retain their form when transformed to any new set of axes at rest referred to the first but differently directed in space, if they are true for any one set they will be true for any other. But we have seen that they are true when $\dot{y}=\dot{z}=0$; they are therefore true in general.

The vectors K_x, K_y, K_z we shall call the components of the force acting in the electron. When q^2/c^2 is negligible the equations of motion are identical with those of Newton; otherwise they are not.

We shall extend the scope of equations (16) so as to embrace the case in which the forces are of gravitational origin. As their applicability in this case is a sheer assumption they can only be regarded as a definition of force.

Energy and Momentum.

If we multiply each component of the universal equations (III) and (IV), Chap. XIII, p. 286, in turn by $E_x, E_y, \ldots H_z$, add them together and integrate over a space at whose boundaries the electric and magnetic forces vanish, we obtain

$$\int\rho\,(E_x u_1+E_y v_1+E_z w_1)\,d\tau+\frac{dL}{dt}=0,$$

where $$L=\int\tfrac{1}{2}\left[(E_x^2+E_y^2+E_z^2)+\tfrac{1}{2}(H_x^2+H_y^2+H_z^2)\right]d\tau$$

is the electromagnetic energy of the space considered. If the electric density ρ is due to electrons of charge e, the integral is equivalent to $\Sigma\, e\,(E_x\dot{x}+E_y\dot{y}+E_z\dot{z})$, where E_x, E_y, E_z are the components of the external electric force (i.e. the part not arising experiments of Kaufmann and Bucherer. As the results of these experiments support Lorentz's calculations they are equally in favour of the relativity theory.

So far it has not been possible to test the relation between V and q experimentally. With the radium rays used by Kaufmann V is not under the control of the experimenter; and with cathode rays it is not possible to get values of q high enough to make the difference between formula (21) and the usual formula $Ve=\frac{1}{2}\mu q^2$ capable of being measured.

The Inertia of Energy.

Consider a physical system surrounded by an imaginary non-material enclosure which is impervious to radiation. The object of this is to prevent the energy of the system from escaping in the radiant form. Let external electric forces whose components are X_e, Y_e, Z_e act on the system. These enable any desired amount of energy to be introduced into the system from outside. The energy gained by the system when referred to the system of reference S, in accordance with our previous results, will be given by the expression

$$\int dE=\int dt\int\rho\,(X_e u_1+Y_e v_1+Z_e w_1)\,d\tau,$$

where ρ is the density of the electricity at any point in the enclosed space. Let us transform this equation so that the right-hand side is referred to variables proper to the system S'. We have

$$t=\beta\left(t'+\frac{v}{c^2}x'\right),\quad x=\beta\,(x'+vt'),\quad y=y',\quad z=z',$$

and

$$t'=\beta\left(t-\frac{v}{c^2}x\right),\quad x'=\beta\,(x-vt).$$

Now,

$$u_1=\frac{u'+v}{1+\frac{u'v}{c^2}},\qquad v_1=\frac{v'}{\beta\left(1+\frac{u'v}{c^2}\right)},\qquad w_1=\frac{w'}{\beta\left(1+\frac{u'v}{c^2}\right)},$$

and

$$\rho=\frac{\rho'}{\beta\left(1-\frac{vu_1}{c^2}\right)}=\beta\left(1+\frac{u'v}{c^2}\right)\rho'.$$

Also

$$X=X',\qquad Y=\beta\left(Y'+\frac{v}{c}N'\right),\qquad Z=\beta\left(Z'-\frac{v}{c}M'\right).$$

So that

$$\int dt \int \rho\,(X_e u_1 + Y_e v_1 + Z_e w_1)\,d\tau$$

$$= \int dt \int dx \int dy \int dz\, F(x, y, z, t)$$

$$= \int dt \int dx \iint dy' dz'\, F'(x', y', z', t')$$

$$= \frac{1}{\beta} \int dt \iiint d\tau'\, F'(x', y', z', t')^*$$

$$= \beta \int dt' \int \rho'\,(X_e' u' + Y_e' v' + Z_e' w')\,d\tau'$$

$$+ \beta v \int dt' \int \rho' \left(X_e' + \frac{v'}{c} N_e' - \frac{w'}{c} M_e'\right) d\tau'.$$

Since the principle of relativity must also apply to the system S' this may be written

$$\int dE = \beta \int dE' + \beta v \int [\Sigma K_x']\,dt' \quad \text{..............(22).}$$

Consider the case in which the motion of the system as a whole is such that it is at rest relative to S', and suppose further that the velocities of its parts relative to S' are so small that v^2/c^2 may be considered negligible. The centre of mass of the system is thus at rest relative to S' which, under the further condition postulated, can only be the case provided $\Sigma K_x' = 0$ for all values of t'. In spite of this $\int[\Sigma K_x']\,dt'$ will not necessarily vanish; for this integral is not taken between given values of t' but of t, so that in general the limits will involve x' as well as t'.

If, however, the external forces do not act except during the interval considered, the parts which would otherwise depend on x' also vanish. This statement will be made clearer in the sequel where the general case is considered. If the forces vanish entirely outside the time limits then we have

$$\int dE = \beta \int dE',$$

so that

$$dE = \beta\,dE'.$$

We therefore conclude that the energy of a uniformly moving system which is not under the influence of external forces is a function of two variables, namely its energy E_0 relative to a system supposing that weight as well as mass is proportional to $\mu + E/c^2$. In the case of radioactive change the matter is more hopeful. The decrease of mass of a system due to the loss of 1000 gm. calories is $4{\cdot}6 \times 10^{-11}$ gm. Now 1 gm. atom of radium in radioactive equilibrium evolves about $3{\cdot}024 \times 10^4$ gm. cals. per hour. Thus its diminution of mass per hour would be

---

* See Williamson's *Integral Calculus*, p. 320.

$$\frac{3{\cdot}024 \times 10^4 \times 4{\cdot}2 \times 10^7}{9 \times 10^{20}} = 1{\cdot}4 \times 10^{-6} \text{ mgm.}$$

This would amount in a year to ·012 mgm. or in 100 years to 1·2 mgm. It would be worth the while of anyone who could afford the necessary capital to make observations of this character.

Momentum.

We have seen that ξ, the x component of momentum of a point charge, satisfies the relation $\frac{d\xi}{dt} = K_x$. Let us apply this result to the system, previously considered, which is surrounded by a closed boundary impervious to radiation and is subjected to the action of external electric forces X_e, Y_e, Z_e, etc. Then the total momentum gained by the system in the time during which the forces act will be

$$\begin{aligned}\int d\xi = \int K_x dt &= \int dt \int \rho \left(X_e + \frac{v_1}{c} N_e - \frac{w_1}{c} M_e \right) d\tau \\ &= \beta \int dt' \int \rho' \left(X_e' + \frac{v'}{c} N_e' - \frac{w'}{c} M_e' \right) d\tau' \\ &\qquad + \frac{\beta v}{c^2} \int dt' \int \rho' \left(X_e' u' + Y_e' v' + Z_e' w' \right) d\tau' \\ &= \beta \frac{v}{c^2} \int dE' + \beta \int [\Sigma K_x'] \, dt'.\end{aligned}$$

Let the system move as before, so that its centre of gravity remains at rest referred to S', then $\Sigma K_x' = 0$, and if in addition the forces are zero outside the time limits considered, the second integral vanishes and therefore

$$\int d\xi = \beta \frac{v}{c^2} \int dE'$$

and

$$d\xi = \beta \frac{v}{c^2} dE' \text{ ....(24).}$$

The Principle of Relativity and the Aether.

Before leaving this part of our subject it is desirable to review the bearing of the principle of relativity on the question of the existence of the luminiferous aether. We have seen that if we start from the hypothesis that electromagnetic actions have their seat in a medium, the aether, which is at absolute rest, the known facts can be explained on the hypothesis of Fitzgerald that bodies contract when in motion relative to the aether. This contraction can be shown to be a plausible consequence of the motion through the aether. If this contraction is the only change due to the motion, the effects in moving systems would not be exactly correlated with those in fixed systems, although the differences, so far as the writer is able to judge, would not have been detected in any experiments which have been carried out up to the present. Any effects which might arise which were not accounted for by such a scheme might be explained by making the velocity of light a function of the motion of the system through the aether.

On the other hand, the principle of relativity, which is in accordance with all the known facts, describes them in a simpler and more symmetrical manner. It is clear that if the principle of relativity and its consequences are valid, electromagnetic experiments can never yield any information as to the state of rest or uniform motion of an aether. This follows since, in the last analysis, all the effects are then made to depend on the relative motion of matter. It is, in fact, quite unnecessary ever to bring the word aether into the discussion.

From this standpoint it is desirable, perhaps, to state the matter somewhat more explicitly. To specify any physical event it is necessary to locate it in time and space, that is to say, to determine the four coordinates t, x, y and z of the time and place at which it occurs. The question arises as to whether we can be sure that two events which appear to occur at the same place at successive times t and t' really occur at the same place. Can we be sure that the place which appears to be the same has not changed its position in the interval? It is clear that such a discussion is futile until we have fixed on a set of axes Ox, Oy, Oz in terms of which we can specify the position of the points considered.

Let us suppose that we have fixed upon such a set of axes and that they are so chosen that the physical system in which the events occur is at rest as a whole when referred to these axes. Having marked off our axes in terms of measuring rods, which we may suppose to be part of the system, we should naturally turn our attention to the measurement of time. This could be done by means of a series of clocks which we could compare with one another at some fixed point, let us say at the origin. It would, however, be necessary to have some means of comparing them when they were moved away so as to be at a considerable distance from one another. This could be done by sending light signals. The simplest assumption we could make would be that the light was propagated in spherical waves. By considering the case of a wave propagated from the origin we see that x, y, z and t would satisfy the relation $x^2+y^2+z^2-c^2t^2=0$, where c is the velocity of light in space. The coordinates as thus determined would be consistent with each other and would be the simplest ones in terms of which the events in that system could be described. If we had no opportunity to investigate other systems we should probably conclude that our system was at rest relative to the medium in which the light was propagated.

Now suppose that we have another system, let us say a distant solar system, which is moving with the velocity v relative to the first. An investigator located in the second system would be able to discover a framework of axes and a set of coordinates x_1, y_1, z_1 and t_1 in terms of which light would in his system be propagated in spherical waves. He would find these coordinates the simplest in terms of which he could describe the events occurring in his own system, and he would have the same reason for concluding that the second system was at rest relative to the aether that the first observer had had for concluding that his was. It is evident that one of the two conclusions must be fallacious, and there is nothing to favour one rather than the other.

The coordinates x, y, z, t and x_1, y_1, z_1, t_1 which refer to a given event are of course different for the two systems. Each set is preferable to the other for describing events in its own system; but if we are to describe events in a universe in which both systems exist there is nothing to choose between them. It is clear that there are an infinite number of possible systems of reference corresponding to all the possible values of v, and each one is as good as another. In fact, if we wish to represent the whole universe in the simplest and most elegant manner, we cannot thus arbitrarily separate time and space, but we must rather consider the whole as a four-dimensional manifold of x, y, z, and t.

Four-dimensional analysis has now been devised with the object of effecting such a representation. It is found that the electromagnetic equations then assume a more symmetrical form than that in which we have considered them. The reader who is interested in these questions may be referred to the following authorities:—

Minkowski, *Raum und Zeit.* Leipzig, 1909.
Sommerfeld, *Ann. der Physik*, vol. XXXII. p. 749, and vol. XXXIII. p. 649. 1910.
M. Laue, *Das Relativitäts Princip.*

To sum up we may say that the Principle of Relativity furnishes no evidence either for or against the existence of an aether. It denies the possibility of determining the motion of such a fluid if it exists. In so far as Relativity is a Universal Principle it finds the aether a superfluous hypothesis.

CHAPTER XV

RADIATION AND TEMPERATURE

It is a very familiar fact that when material bodies are heated they emit electromagnetic radiations, in the form of thermal, luminous and actinic rays, in appreciable quantities. Such an effect is a natural consequence of the electron and kinetic theories of matter. On the kinetic theory, temperature is a measure of the violence of the motion of the ultimate particles; and we have seen that, on the electron theory, electromagnetic radiation is a consequence of their acceleration. The calculation of this emission from the standpoint of the electron theory alone is a very complex problem which takes us deeply into the structure of matter and which has probably not yet been satisfactorily resolved. Fortunately we can find out a great deal about these phenomena by the application of general principles like the conservation of energy and the second law of thermodynamics without considering special assumptions about the ultimate constitution of matter. It is to be borne in mind that the emission under consideration occurs at all temperatures although it is more marked the higher the temperature.

The problem which we set before us is that of finding the nature of the radiation which is found in an enclosure containing material bodies and maintained at a constant temperature. Sufficient time is supposed to have elapsed for any secular changes in the enclosed matter to have come to an end, so that the nature and condition of any element of the matter does not vary. Special radiations of chemical or radioactive origin are therefore eliminated in so far as they involve progressive material changes. Even in the steady state the interchange of radiation will be accompanied by an interchange of electrons arising from thermionic and photoelectric emission (see Chap. XVIII). The effect of this on the frequency ν, the angle θ, and the plane of polarization as well as on the temperature. ϵ_1 and ϵ_2 must be independent of i_1 and i_2 provided the enclosure contains radiating matter other than the body S and provided the area of S is vanishingly small compared with that of the other matter.

The only way in which the body S can either gain or lose energy is by interchange of radiation, so that the net rate at which its energy increases is

$$\int_S \int_0^\infty \int_0^{2\pi} \{i_1(1 - A_1) - \epsilon_1 + i_2(1 - A_2) - \epsilon_2\} \cos\theta \, dS d\nu d\omega \quad \ldots(4).$$

The value of this integral must always be zero, otherwise the temperature of S would alter. This would contravene the second law of thermodynamics; since the difference of temperature thus established could be used to furnish available work, which would then be obtained from a source at a constant temperature.

This conclusion must be true whatever the shape, size and nature of the body S, whatever its position in the enclosure and whatever the nature, shape and size of the enclosure and the other matter contained in it may be; provided only that the enclosure is maintained at a constant temperature.

The fact that the integral (4) vanishes under all these conditions enables us to establish many important properties of the functions i, A and ϵ. In the first place the radiation must be isotropic; that is to say, i must have the same value at a given point for all directions in space. It cannot be a function of ω. If it were, the intensity of the radiation in some directions would be stronger than in others. By taking S to be a flat object it could be turned so as to receive more or less of the stronger radiation in different positions. Differences of temperature would thus be set up which would contravene the second law. This result must be true for all frequencies and all planes of polarization. This is clear since S may be a substance which absorbs some frequencies and transmits others, or it may be a plate of a material like tourmaline which has much more intense absorption for light polarized in certain planes than in others.

By expressing the element of solid angle $d\omega$ in terms of the was discovered by Balfour Stewart, is usually known as Kirchhoff's Law.

The foregoing results require amplification when the nature of the medium varies from one point of the enclosure to another. The relation $i=\epsilon$ no longer proves that i has the same value everywhere in the enclosure, since the emissivity ϵ of a black body may depend, and in fact does depend, on the nature of the surrounding medium. If we call I the function which has the same value at all points in the enclosure it is clear that the relation between i and I must be determined by properties of the medium which have nothing directly to do with absorption or emission. For if we consider a portion of the medium where there is no emitting or absorbing matter, but which is characterized by a particular value of the velocity of transmission, the appropriate value of i will somehow have to be established. In order to find I it is not therefore necessary to concern ourselves directly with the way in which ϵ may be modified according to the nature of the medium in which the emitting system is embedded. All that is necessary is to consider the passage of radiation across the interface between two portions of the medium characterized by different velocities of transmission.

Let us imagine that one of the regions in question is separated from the other by a perfectly reflecting interface. The introduction of this cannot make any difference to the nature of the radiation. The interface is punctured at the point A, leaving a small opening of area dS. In the upper medium, where the velocity of transmission is V, is a perfectly reflecting hemisphere with equal and symmetrical apertures at B and C which subtend equal infinitesimal solid angles $d\omega$ at A. If AD is the direction of the refracted ray corresponding to the incident ray BA, it is clear that with this arrangement the only radiation which can get from the lower medium to the part of the upper medium outside the hemisphere is that comprised in a small solid angle $d\omega'$ about AD and the equal beam which is symmetrical about the normal AF. The rest is all reflected by the hemisphere and returned through dS.

The ratio of the two elements of solid angle is

$$d\omega'/d\omega=\sin\theta'\,d\theta'/\sin\theta\,d\theta,$$

Since L, i and ϵ may be split up into their spectral components it follows that each of these will have to satisfy equations of the type

$$\int_0^\infty (L_\nu, i_\nu, \epsilon_\nu)\, d\nu = \text{const.} \times T^4 \quad\ldots\ldots\ldots\ldots\ldots(11).$$

Reflection of Radiation at a Moving Mirror.

There is another effect, in addition to those already contemplated, which is produced when light or radiation is reflected at a moving surface. This results in a change in the wave-length or frequency of the light, which is akin to Doppler's effect. By considering the simple case in which the mirror is moving parallel to the direction of propagation of the radiation, which is incident normally, it is clear that the frequency of the light is diminished after reflection by an amount which is proportional to the velocity of the reflector, when this is in the same sense as that of the incident radiation. If the reflector is moving towards the incident beam the frequency of the reflected beam is greater than that of the incident beam. The complete resolution of more complex cases is effected most easily by means of the principle of relativity.

We shall consider the case of a plane reflector moving with uniform velocity v parallel to the direction of its normal, which we shall take as the axis of x. In the case of a ray incident at an angle θ the vectors which specify it will only contain the space and time coordinates through the factor

$$\cos n\left(t + \frac{x\cos\theta + z\sin\theta}{c} + \gamma\right),$$

if the plane of incidence is that of the x and z axes and γ is a constant which specifies the phase. The corresponding factor for the reflected ray is

$$\cos n\left(t - \frac{x\cos\theta - z\sin\theta}{c} + \gamma\right),$$

provided the reflector is at rest. If the reflector is in motion these expressions must be unchanged when all the quantities are measured in reference to axes x', y', z' and t' referred to which the

thus
$$\frac{\nu_2}{\nu_1}=\frac{1-v^2/c^2}{1-2\frac{v}{c}\cos\theta+\frac{v^2}{c^2}}=1+2\frac{v}{c}\cos\theta\ \ldots\ldots\ldots(12),$$

neglecting $(v/c)^2$ and higher powers.

The corresponding relation between the wave-lengths is

$$\frac{\lambda_2}{\lambda_1}=1-2\frac{v}{c}\cos\theta\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots(13).$$

Wien's Law.

The foregoing considerations, coupled with thermodynamical principles, enable us to take another step forward towards the discovery of the function $F(\nu, T)$. This advance is due to Wien[*]. The argument resolves itself into two parts. In the first place, if we start with an enclosure containing only thermal radiation, such as we have seen to be characteristic of some temperature T, and then alter the nature of this radiation by means of a motion of some part of the perfectly reflecting boundary wall, we shall be able to show that the resulting radiation is invariably such as is characteristic of some other undetermined temperature T'. Having established this proposition, the second step consists in making use of it so as to find out as much as we can about the nature of the function $F(\nu, T)$.

Consider the cylinder with perfectly reflecting walls shown in Fig. 41. The ends are closed by plates of radiating matter maintained at the temperatures T_1 and T_2 respectively. T_2 is greater than T_1. The transverse partitions are perfectly reflecting and are provided with shutters D_1 and D_2 which can be opened or closed at will. The partitions can also be caused to slide along the cylinder. We now imagine the following processes to occur:—(1) D_2 is shut and D_1 open. Then C is filled with radiation characteristic of T_2 and A and B with radiation characteristic of T_1. When equilibrium has become established D_1 is closed. (2) The radiation in C is allowed to push the piston F and so compress the radiation in B until the pressure of the radiation in B is equal to that in C. On account of the Doppler effect the nature of the radiation in B will have become changed temperature T. The face of the piston is perfectly smooth and perfectly reflecting. The walls of the cylinder reflect completely, i.e. without loss of energy, but irregularly in all directions. This device keeps the radiation isotropic and so makes it unnecessary for us to consider certain complications which would otherwise arise. The sectional area of the cylinder is A and its height h. h will vary with the position of the piston. Let $L(\lambda)\,d\lambda$ denote the energy per unit volume of the cylinder which belongs to wave-lengths between λ and $\lambda + d\lambda$, then $J = Ah\,L(\lambda)\,d\lambda$ is the total energy belonging to these wave-lengths which is present in the cylinder. Owing to the motion of the piston the value of J will tend to change. It will tend to decrease because the radiation whose wave-length is near λ will join some other group when it is reflected at the moving piston, and it will tend to increase because other radiation will have its radiation changed to values near λ when it is reflected. The rate of loss of energy by the group is clearly equal to the total amount which is incident on the moving piston in unit time. The proportion of J which belongs to rays whose direction of propagation lies within any small solid angle $d\omega = 2\pi \sin\theta\, d\theta$ is $d\omega/4\pi$, since the radiation is isotropic. Thus the amount of energy belonging to these wave-lengths which is incident on the piston in unit time in directions lying within the cones whose semi-angles about the normal are θ and $\theta + d\theta$ is

---

[*] *Berl. Ber.* 9 Feb. 1893.

$$\tfrac{1}{2} L(\lambda)\,d\lambda \,.\, \sin\theta\, d\theta \,.\, cA \cos\theta.$$

The rate of loss of energy by the group will be obtained by integrating this expression over all angles θ which lie between 0 and $\pi/2$. It is thus equal to

$$\tfrac{1}{4} Ac\,L(\lambda)\,d\lambda \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(14).$$

The calculation of the rate of gain of energy by the group is more complicated. If we consider a ray of wave-length λ' incident at an angle θ the wave-length of the reflected ray will be given by

$$\lambda' = \lambda\left(1 - \frac{2v}{c}\cos\theta\right) \text{ by (13).}$$

The change of sign is due to the fact that we are now taking v to be positive for an outward motion of the piston, which corresponds to increasing volume. Thus, by reflection, rays

But by Stefan's law $L = A_1 T^4$, where A_1 is a universal constant. Thus

$$4A_1 T^3 \frac{\partial T}{\partial t} = \frac{\partial L}{\partial t} = -4\kappa A_1 T^4$$

or

$$-\kappa\, dt = \frac{dT}{T};$$

so that (21) may be written

$$T\frac{\partial L(\lambda)}{\partial T} = 5L(\lambda) + \lambda\frac{\partial L(\lambda)}{\partial \lambda} \quad \text{...............(22).}$$

If the variables are changed to $T_1 = \log T$, $\lambda_1 = \log \lambda$ and $L_1 = \log L(\lambda)$, (22) becomes

$$\frac{\partial L_1}{\partial T_1} = 5 + \frac{\partial L_1}{\partial \lambda_1}.$$

Now change the independent variables from T_1 and λ_1 to T_1 and π_1, where $\pi_1 = T_1 + \lambda_1$. Denoting the new partial derivatives by δ, we have

$$\frac{\partial L_1}{\partial \lambda_1} = \frac{\delta L_1}{\delta \pi_1}\frac{\delta \pi_1}{\delta \lambda_1} + \frac{\delta L_1}{\delta T_1}\frac{\delta T_1}{\delta \lambda_1} = \frac{\delta L_1}{\delta \pi_1},$$

and

$$\frac{\partial L_1}{\partial T_1} = \frac{\delta L_1}{\delta \pi_1}\frac{\delta \pi_1}{\delta T_1} + \frac{\delta L_1}{\delta T_1}\frac{\delta T_1}{\delta T_1} = \frac{\delta L_1}{\delta \pi_1} + \frac{\delta L_1}{\delta T_1},$$

so that

$$\frac{\delta L_1}{\delta T_1} = 5,$$

and

$$L_1 = 5T_1 + F(\pi_1),$$

where $F(\pi_1)$ is an arbitrary function of π_1. Substituting for the logarithms and multiplying both sides by $d\lambda$ we find that

$$L(\lambda)\, d\lambda = T^5 \phi(T\lambda)\, d\lambda \quad \text{........................(23)}$$

$$= \lambda^{-5} (T\lambda)^5 \phi(T\lambda)\, d\lambda$$

$$= \lambda^{-5} \chi(T\lambda)\, d\lambda \quad \text{....................(24),}$$

where ϕ and χ are undetermined functions of the product $T\lambda$ only.

It is known from experiment that $L(\lambda)$ has the value zero when $\lambda = 0$ or ∞ and a single maximum between these limits. Differentiating (24) by λ we see that the maximum and minimum values of $L(\lambda)$ satisfy the equation

$$\frac{\partial L(\lambda)}{\partial \lambda} = \lambda^{-6}\left\{\lambda T \frac{\partial[\chi(\lambda T)]}{\partial(\lambda T)} - 5\chi(\lambda T)\right\} = 0.$$

where k_1, k_2, and k_3 are integers. Since $l^2 + m^2 + n^2 = 1$, we have

$$\frac{k_1^2}{d_1^2} + \frac{k_2^2}{d_2^2} + \frac{k_3^2}{d_3^2} = \frac{4}{\lambda^2}.$$

Thus k_1, k_2, and k_3 can have any integral values which satisfy this equation. Since the equation is that of an ellipsoid whose coordinates are k_1, k_2, and k_3, and whose semi-axes are $2d_1/\lambda$, $2d_2/\lambda$, and $2d_3/\lambda$, the number of vibrations whose wave-lengths lie between λ and $\lambda + d\lambda$ is equal to the number of points having coordinates which are positive integers which lie in the volume between the ellipsoids whose parameters are determined by λ and $\lambda + d\lambda$ respectively. The number of such points is equal to the volume lying in a single octant between the two ellipsoidal shells. The number is therefore $4\pi d_1 d_2 d_3 d\lambda/\lambda^4$ and the energy per unit volume of the box and wave-lengths between the limits λ and $\lambda + d\lambda$ is $8\pi RT d\lambda/\lambda^4$. It has been found that every purely dynamical method of calculating the radiation formula leads to (26) and (27).

Although the value of $L(\lambda)$ given by (26) satisfies Wien's functional relation (24) it is quite inconsistent with the results of experiments on the complete temperature radiation. The experiments show that $L(\lambda)$ has a maximum for a value of λ whose position is governed by Wien's displacement law and has the value zero when $\lambda = 0$. But (26) would make $L(\lambda)$ increase indefinitely as λ approached zero. There are a number of ways in which this contradiction might conceivably be avoided. Jeans has suggested that the final state of equilibrium, if it could ever be attained, would be given by (26); but that the rate of emission of the energy of short wave-lengths by matter is so slow that the distribution given by (26) would only become established after an indefinite lapse of time, and in any finite time would not even be approximately realized. According to this view the observed distribution is one of intermediate equilibrium due to less fundamental causes. Nevertheless the causes which establish the intermediate distribution must be sufficiently fundamental to be common to all types of matter; otherwise the degree of consistency which has been observed in experiments on thermal radiation would not be found.

Another way of escape is to deny the applicability of the theorem of equipartition of energy to aetherial, as opposed to material, systems. There are a number of reasons why this is not a very satisfactory alternative. In the first place the same kind of objection may be made against the applicability of equipartition to material systems. Thus the specific heats of gases are not what would be given by a naïve application of this theorem, particularly when one considers the large number of electrons present in the atoms and that each electron ought to have its quotum of energy. On the other hand the law of equipartition has been so successful in other directions that it is difficult to believe that its deduction from dynamical principles involves a fundamental error. In the second place the dilemma in question does not seem to be connected with the law of equipartition alone, so that denial of the validity of this would not really remove the difficulty. It appears that a number of calculations which use the first method, that of equilibrating the absorption and emission of energy, lead to Rayleigh's formula (see for example Chap. XVII, p. 433). In fact this formula appears inevitably to arise whenever the emission and absorption of radiant energy by matter is assumed to be a continuous process subject to dynamical and electrodynamic laws.

Although it may appear very revolutionary to some, it seems to the writer that the only logical way out of these difficulties is to deny the adequacy of dynamics and electrodynamics for the explanation of the emission or absorption of radiation by matter. If we leave the familiar guidance of dynamics behind it is necessary to find some other set of fundamental principles to rely on. Although the matter cannot yet be regarded as fully and satisfactorily worked out, a valuable start in this direction has unquestionably been made by Planck and his followers.

Planck's Law.

We have stated that Planck has shown how theoretically to deduce a radiation formula which is in good agreement with the results of experiment. It would take up too much space adequately to consider the various vicissitudes which this theory has undergone at the hands of Planck and his critics, so that we shall practically confine ourselves to the discussion of the most recent form* of it.

---

* *Ann. der Phys.* vol. XXXVII. p. 642 (1912).

Thus the mean energy U of the oscillators, and also, since

$$\frac{1}{\eta} = 1 + pL(\nu_0),$$

the division of the energy among them in the steady state, is completely determined.

In order to introduce the temperature T it is necessary to calculate the entropy S of the system. This is equal (see Chap. XVII, pp. 400 and 407) to R times the logarithm of the probability of the system, defined in a similar way to that introduced by Boltzmann* into the kinetic theory of gases. The present case is a little different from that contemplated in the kinetic theory, inasmuch as it is only the part of the energy of the resonators which is an integral multiple of ϵ that is a matter of chance and therefore subject to probability considerations†. The probability sought is the number of ways in which the N resonators can be arranged so as to have the given distribution of energy units, subject to the condition that the same distribution arises whenever the resonators which have a given number of units are interchanged among themselves. This probability may be calculated as follows:

Along the positive energy axis lay off marks at the points $0, 1\epsilon, 2\epsilon \ldots m\epsilon \ldots \infty\,\epsilon$. From these marks draw lines perpendicular to the axis and on them lay off $NP_0, NP_1, NP_2 \ldots NP_m \ldots NP_\infty$ equidistant dots. Altogether there are $\Sigma NP_m = N$ dots and their distribution over the diagram is a geometrical representation of the way in which the resonators are distributed about the energy of the system. Since the dots which have equal numbers of units are considered to give rise to systems which are indistinguishable, the number of independent ways in which the N dots can be arranged to form the given distribution is

$$W = N! \div \Pi\,(NP_m)!$$

where Π denotes the continued product for all the values of m from 0 to ∞. Thus

$$S = R \log W = R \log N! - R\Sigma \log (NP_m)!,$$

---

* *Vorlesungen über Gastheorie*, p. 41, Leipzig (1896).

† Planck, *Sitzber. der K. Pr. Akad. der Wiss.* vol. XXXV. p. 723 (1911).

where $x=\frac{ch}{R\lambda T}$. (51) has two real roots, viz. $x=0$ and $x=4{\cdot}9651$. Thus the maximum value is given by substituting $x=4{\cdot}9651$ in

$$\lambda_m T=\frac{ch}{Rx}=b \quad\text{.........................(52).}$$

According to the measurements of Lummer and Pringsheim* $b=0{\cdot}294$ cm. deg. Solving (50) and (52) for R and h we find

$$\left.\begin{aligned}R&=1{\cdot}346\times10^{-16}\text{ erg deg.}^{-1}\\h&=6{\cdot}548\times10^{-27}\text{ erg sec.}\end{aligned}\right\}\quad\text{...............(53).}$$

Since R is the constant in the gas equation $pv=R_1T$ reckoned for a single molecule, this value of R enables us to deduce the value of N, the number of molecules in a cubic centimetre of a gas under standard conditions, and the charge e on an ion, from well-known data. We have $pv=NRT$, where

$$p=76\times13{\cdot}6\times981\text{ dynes cm.}^{-2},\quad v=1\text{ cm.}^3\quad\text{and}\quad T=273\text{ deg.}$$

Whence $$N=2{\cdot}76\times10^{19}\text{ per cm.}^3.$$

Since the charge which is required to liberate half a cubic centimetre of hydrogen, measured under standard conditions, in electrolysis is 0·4327 E.M. units (see Chap. I, p. 6) it follows that

$$Ne=0{\cdot}4327\text{ E.M.U.},$$

whence $$e=4{\cdot}69\times10^{-10}\text{ electrostatic units}$$

$$=18{\cdot}9\times10^{-10}\text{ in our units.}$$

These values of N and e are in excellent agreement with those which have been found by Millikan, Rutherford and others, using more direct methods. This agreement must be regarded as supporting very strongly those assumptions in Planck's demonstration which are necessary to produce the formulae finally obtained.

We shall see that Planck's radiation theory has recently received unexpected support in two other directions. One of these will now be considered.

Radiation and Specific Heat.

From the phenomena exhibited by absorption bands, the residual rays, and so on, we know that something like Planck's oscillators must exist in actual matter and possess natural where α is the coefficient of linear expansion, V is the atomic volume, and K the compressibility. The difference between C_p and C_v is only appreciable at the higher temperatures.

---

* *Verhand. der Deutsch. Physik. Ges.* vol. I. p. 230.

It will be observed that (55) contains only one adjustable constant, the frequency ν, so that the agreement shown by the table is quite convincing. In fact in the case of KCl, KBr and NaCl the value of ν has been taken from the experiments of Rubens and Hollnagel on the residual rays from these substances. Thus in these cases the complete thermal behaviour may be predicted from the determination of a single optical constant, and shows an excellent agreement with the observed results. In other cases, for instance mercurous chloride, it is necessary to use a formula involving a summation over two different values of ν; but there are good reasons for assigning at least two frequencies for compound substances, so that this cannot be considered to be an argument against the general position.

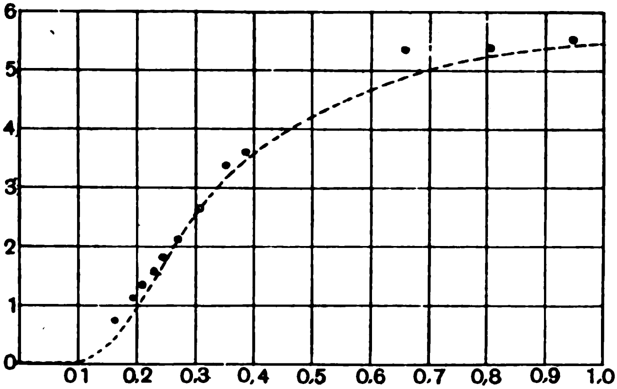


Fig. 42.

The general character of the relation between C_v and $\frac{h\nu}{RT}$, given by (54) and (55), is shown by the accompanying diagram (Fig. 42) taken from Einstein's paper. The ordinates represent the values of the right-hand side of (54) and the abscissae those of $\frac{h\nu}{RT}$. NR is taken $= 5{\cdot}94$. The calculated values of C_v are shown by the broken curve. The circles represent the older observations of the specific heat of carbon, an appropriate value in different portions of matter in the way in which the electric charges can be separated. This would be extremely unlikely to be the case unless the fundamental magnetic element contained both positive and negative magnetism and were therefore similar to an electric doublet rather than to the positive and negative charges of which such a doublet is constituted.

This basic difference receives a ready explanation on the electron theory. According to that theory magnetic forces can arise only from the motion of electric charges and, in the last analysis, from the motion of electrons. Now there is no possible motion of electrons which can give rise to an isolated magnetic pole; but there is a very simple type of motion which gives rise to a system which is the magnetic analogue of the electric doublet, i.e. a system which has positive magnetism on one side of it and negative on the other. Many years ago Ampère built up a theory of magnetic media on the assumption that the atoms were the seat of circular electric currents. As is well known, such a current behaves like a small magnet, and the hypothesis is therefore all that is required to account for magnetic polarization and hence, from the analogy with dielectrics, for magnetizable media. Now we shall be able to show that an electron revolving in a closed orbit is equivalent to a small magnet in the same way that Ampère's atomic currents were. This theory, whose development is due largely to Weber and Langevin, will be shown to give a simple explanation of diamagnetism as well as paramagnetism. With certain further assumptions which do not seem improbable, it can be made to give a good account of the more complex phenomena of ferromagnetism as well.

The Magnetic Force due to a Moving Electron.

We have seen (Chap. XI, p. 221) that the components of the magnetic intensity due to an electrically charged particle moving with the uniform velocity w parallel to the z axis are

$$H_x = \frac{\beta}{\sqrt{1-\beta^2}} \frac{\partial \phi_1}{\partial y_1}, \quad H_y = -\frac{\beta}{\sqrt{1-\beta^2}} \frac{\partial \phi_1}{\partial x_1}, \quad H_z = 0,$$

where $\beta = w/c$. When β is small, which we shall suppose to be the case in the motions we are dealing with, we can put

$$\phi_1 = e/r, \quad x_1 = x \quad \text{and} \quad y_1 = y.$$

orbit the average force in the y direction at P arising from the $\dot{x}$ motions at Q and R is

$$\frac{e\dot{x}}{2c}\left\{\frac{1}{(OP-z)^2}-\frac{1}{(OP+z)^2}\right\}=\frac{e\dot{x}}{2cOP^2}\left\{\left(1-\frac{z}{OP}\right)^{-2}-\left(1+\frac{z}{OP}\right)^{-2}\right\}$$

$$=\frac{e\dot{x}}{c\,.\,OP^2}\times\frac{2z}{OP}.$$

The average value of this taken all round the orbit is

$$2\frac{e}{c\,.\,OP^3}\frac{1}{\tau}\int z\frac{dx}{dt}\,dt=2\frac{e}{c\,.\,r^3}\frac{\text{area}\,SV'QR'}{\tau}=2\frac{e}{c\,.\,r^3}\frac{S}{\tau}\sin\theta.$$

The sign of this force is determined by the direction of the motion at the side of the orbit nearest to P and is evidently opposite to that arising from the component $\dot{z}$. The balance of tangential force is thus

$$H_t=\frac{e}{c\,.\,r^3}\frac{S}{\tau}\sin\theta \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5).$$

The average value of the component of the magnetic force perpendicular to both OP and PT vanishes. Because if AB is the line of intersection of the plane NOP with the plane of the orbit the latter can be divided into pairs of elements dS_1, dS_2 which are symmetrical about AB and are equidistant from P. These elements produce equal and opposite effects at P so far as the component under consideration is concerned. The orbit can thus be divided into mutually interfering pairs of points so that the average value of this component is zero.

It is evident from formulae (4) and (5) that the average value of the magnetic field of force due to the revolving electron is exactly equivalent to that which would be given by a small magnet whose moment is

$$M=\frac{eS}{c\tau}\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

and whose axis coincides with the axis of revolution of the electron.

An atom may in general contain a number of electrons rotating in closed orbits as well as others which execute small oscillations about a position of static equilibrium. The orbits may be numerous and distributed in various azimuths inside the atom, so as to furnish no resultant magnetic moment; or they may revolving electrons produced by the application of a given external field H. Let S' be the area of the orbit projected on a plane perpendicular to H. Then the establishment of an external field H will cause a flux HS' of magnetic force through the orbit. If the orbit is circular this will give rise to an electric intensity E tangential to the orbit, where

$$\int E ds = -\frac{1}{c}\frac{d}{dt}(HS\cos\theta) = -\frac{\delta(HS\cos\theta)}{c\tau}$$

$$= -\frac{\omega}{2\pi c}\delta(HS\cos\theta) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7),$$

where θ is the angle between H and the normal to S, τ is the periodic time, ω the angular velocity and δ denotes the change per revolution.

But the moment of the force Ee will cause an increase in the moment of momentum of the electron. Thus

$$Eer = \frac{d}{dt}(mr^2\omega) = 2\frac{mc}{e}\frac{dM}{dt},$$

since
$$M = \frac{eS}{c\tau} = \frac{e}{c}\frac{\omega r^2}{2}.$$

So that
$$dM = \frac{e^2}{2mc}Er\,dt = \frac{e^2}{2mc\omega}E\,ds,$$

and
$$\delta M = \frac{e^2}{2mc\omega}\int E\,ds = -\frac{e^2}{4\pi mc^2}\delta(HS\cos\theta),$$

neglecting the change in ω during one revolution. Hence

$$\frac{\delta M}{M} = -\frac{\tau}{4\pi}\frac{e}{mc}\frac{\delta(HS\cos\theta)}{S} \quad \ldots\ldots\ldots\ldots\ldots\ldots (8).$$

If we neglect the change in S compared with S, and if we take $\tau = 10^{-15}$ sec. to correspond with spectral lines, since

$$\frac{e}{m} = 1{\cdot}77 \times 10^7 \times c\sqrt{4\pi},$$

$\delta M/M$ is of the order $10^{-9} \times \delta H$. Since the greatest attainable magnetic fields correspond to $H < 10^5$ it follows that the effect we are considering will not change the magnetic moment of the orbits by more than about 10^{-4} of its value.

We have seen that the effect of applying a magnetic field is to change the periods of the orbits projected into a plane perpendicular to the magnetic field, the component of the motion parallel to the latter being unaltered. Since the revolving electrons are accelerated they will in general be radiating, and the frequencies of the radiation will be determined by the periods of revolution. There will thus be a change in the frequency of the radiation produced by the magnetic field. The displacement of the spectral lines due to this cause was discovered by Zeeman in 1896. The foregoing method is not a satisfactory one for determining quantitatively the changes in the frequency and character of the radiation which arise in this way, as it does not sufficiently consider the motions perpendicular to the orbit. This deficiency will be remedied when the theory of the Zeeman effect is considered in Chapter XX. Without going more deeply into the matter at this stage, it is evident from what has been said, that the magnetic displacement of the spectral lines and the phenomenon of diamagnetism are very intimately related, on the theory we are discussing.

It appears that the occurrence of electrons revolving in orbits is quite unnecessary to account for diamagnetism. The same kind of effects occur even if the electrons are at rest before the magnetic field is applied. This is shown very clearly by the following investigation which is due to Lorentz*.

We shall confine ourselves to arrangements of electrons which are isotropic with respect to three mutually perpendicular directions. Let the coordinates of any particular electron with respect to any set of rectangular axes, whose origin O is the centre of mass of the system, be x, y, z. Then $\Sigma x = \Sigma y = \Sigma z = 0$, taken over the whole body. Let its moment of inertia about any axis through O be $I = 2mk$. Then

$$k = \Sigma x^2 = \Sigma y^2 = \Sigma z^2 \text{ and } \Sigma xy = \Sigma xz = \Sigma yz = 0.$$

Let E be the resultant electromotive intensity of external origin at any point, then the force acting on an electron has the components eE_x, eE_y, eE_z, and the couple about O has the components

$$e\Sigma(yE_z - zE_y),\quad e\Sigma(zE_x - xE_z),\quad e\Sigma(xE_y - yE_x).$$

---

* *Theory of Electrons*, p. 124.

Paramagnetic Gas.

An instructive case, which has been considered by Langevin, is that of a paramagnetic gas such as oxygen. In this case the kinetic theory of gases enables us to calculate the form of the function f in the relation $I=f\left(\frac{H}{T}\right)$ which is required by the laws of thermodynamics. We know from Boltzmann's Theorem (see Chap. XVII, p. 403) that if the molecules of a gas in a closed vessel occupy positions in which they have varying amounts of potential energy, then there will be a greater number per unit volume in the positions in which the potential energy is less. In fact the ratio of the concentrations of the gas at two points where the potential energy differs by w is $e^{-w/RT}$, where R is the gas constant for a single molecule.

In the present instance w will be the potential energy of the equivalent magnet in the field H, i.e. $-MH\cos\theta$, where θ is the angle between the magnetic axis of the molecule whose moment is M and the field H. The number of molecules dn whose magnetic axes lie within the two cones whose semi-angles are θ and $\theta+d\theta$ respectively is

$$dn = 2\pi A e^{\frac{MH}{RT}\cos\theta}\sin\theta d\theta,$$

where A is a constant as yet undetermined. The total number N of molecules considered is evidently

$$N = 2\pi A\int_0^\pi e^{\frac{MH}{RT}\cos\theta}\sin\theta d\theta$$

$$= 2\pi A\frac{RT}{MH}\left\{e^{\frac{MH}{RT}} - e^{-\frac{MH}{RT}}\right\}.$$

Thus $$A = \frac{N}{4\pi}\frac{MH}{RT}\bigg/\sinh\frac{MH}{RT}.$$

The resultant intensity of magnetization I is in the direction of H, by symmetry, and is given by

$$I = \int M\cos\theta\, dn = 2\pi AM\int_{-1}^{+1} xe^{\frac{MH}{RT}x}dx$$

$$= 4\pi AM\left\{\frac{\cosh a}{a} - \frac{\sinh a}{a^2}\right\},$$

where $a = MH/RT$. But $A = Na/4\pi \sinh a$, so that

$$I = MN\left(\frac{\cosh a}{\sinh a} - \frac{1}{a}\right) \quad \text{..................(13).}$$

Since M is determined by the structure of the molecules, we see that for a given density of gas I depends only on a, i.e. I is a function only of H/T in accordance with the conclusion already reached by thermodynamic reasoning. We also note that I is proportional to N, i.e. to the pressure of the gas.

It has not been possible as yet to test this formula by experiments on gases on account of the smallness of the intensities of magnetization which they acquire. It has, however, been extensively used by P. Weiss in building up a theory of the behaviour of ferromagnetic substances, as we shall see.

Ferromagnetic Substances.

The main difference between the ferromagnetic and the paramagnetic substances lies in the very high intensities of magnetization attainable by the former, in the fact that they are capable of permanent magnetization, and that the magnetization is not in general a definite function of the external magnetic force H. The magnetization does not change reversibly with H and substances of this class therefore exhibit the phenomenon of hysteresis. It results that the heat $Q = \int H\,dI$, which is developed when H is made to pass through a cycle of changes, no longer vanishes but has a finite value.

Weiss* has attempted to explain the facts of ferromagnetism on the hypothesis that the only forces which it is necessary to consider as acting on the elementary magnets or revolving electrons are (1) the impressed magnetic field H_1 of external origin, and (2) the molecular field H_2 arising from the elementary magnets of the neighbouring atoms. It is also necessary to take into account the kinetic reactions arising from the thermal agitation of the molecules, just as in Langevin's theory of a paramagnetic gas.

The assumption of a molecular field uniform throughout the substance will naturally be no more accurate than the similar assumptions as to uniformity which are almost invariably made in dealing with molecular physics. This field will vary greatly from point to point but we may reasonably expect to get valuable indications of the way in which a real body would behave by assuming for H_2 a uniform value equal to its average value throughout the substance. The theory entirely omits to consider interatomic forces of non-magnetic type. This is clearly legitimate until it is shown that such forces do play an important part in the phenomena under consideration.

---

* *Journal de Physique*, vol. vi. p. 661 (1907).

Permanent Magnetization.

Weiss's theory accounts for the existence of permanent magnetization in the following fashion. Bearing in mind that the molecular magnets are in equilibrium under the influence of thermal agitation and the intensities H_1 and H_2, the value of I for a given value of H will be determined by the equations

$$\frac{I}{I_0} = \frac{\cosh a}{\sinh a} - \frac{1}{a}, \quad a = \frac{MH}{RT} \text{ and } H = H_1 + H_2,$$

where $I_0 = MN$ is the maximum possible intensity of magnetization which is attained when the axes of the elementary magnets all point in the same direction. Moreover H_2 is proportional to the intensity of magnetization and may be written $H_2 = \lambda I$, where λ is some constant. Permanent magnetization corresponds to the absence of external field, so that $H_1 = 0$, and we have the two independent equations

$$\frac{I}{I_0} = \frac{\cosh a}{\sinh a} - \frac{1}{a} \text{ and } \frac{I}{I_0} = \frac{RT}{\lambda M I_0} a \text{............} (14).$$

The values of I which satisfy these equations may be found most readily by a graphical method (see Fig. 47). Let ORP be the graph of $\frac{I}{I_0} = \frac{\cosh a}{\sinh a} - \frac{1}{a}$ and OQP that of $\frac{I}{I_0} = \frac{RT}{\lambda M I_0} a$. According to the simplest hypothesis which can be made R, λ, M and I_0 are constants independent of T and H, so that OQP is a straight line which makes an angle α with the axis of x, where $\tan \alpha = \frac{RT}{M\lambda I_0}$ is proportional to the absolute temperature T. The possible values of I/I_0 are given by the intersections of OQP and ORP. There are thus always two possible values of I, one of which is zero. Of these, however, the point P corresponds to a stable and the point O to an unstable condition of the substance. To show this, let us suppose that by means of an external field the intensity of magnetization is made to undergo a slight decrease, so that it is now determined by the line ST parallel to the axis of a. This state of magnetization gives rise to an internal field which is proportional to OU, whereas to overcome the thermal tendency to disorganization it is only necessary to have a field proportional to OV. The magnetization of the substance will therefore increase automatically until the state corresponding to the point P is reached. The reverse happens if the magnetization is given a virtual increase beyond that which corresponds to the point P. Thus P represents a stable configuration of the material. In the same way it can be shown that O represents an unstable condition.

The permanent magnetization exhibited by the paramagnetic metals is not as definite as this theory would lead one to expect.

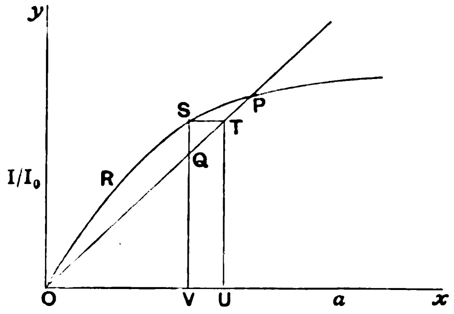


Fig. 47.

The indefiniteness may, however, be due to the fact that these materials are not microscopically homogeneous, as well as to the occurrence of microscopic local reversals of magnetization. There are also, in all probability, complications arising from the crystalline character of these materials.

As the temperature rises the slope of the line OP increases until at a certain temperature T_0 it coincides with the tangent to the curve at the origin. At higher temperatures than this the only possible solution would be $I = 0$, so that the substance would be incapable of permanent magnetization in the absence of an external magnetic field. The temperature T_0 may therefore be interpreted as that at which the ferromagnetic quality disappears.

To determine the slope of the tangent to the curve at the origin we have

$$\frac{d}{da}\left(\frac{I}{I_0}\right)=\frac{d}{da}\left(\frac{\cosh a}{\sinh a}-\frac{1}{a}\right)$$

$$=1-\left(\frac{e^a+e^{-a}}{e^a-e^{-a}}\right)^2+\frac{1}{a^2}$$

$$=1-\frac{1}{a^2}\left(1+\frac{a^2}{2}\right)^2\left(1+\frac{a^2}{6}\right)^{-2}+\frac{1}{a^2}$$

$$=\tfrac{1}{3} \text{ when } a=0.$$

Thus $$\frac{1}{3}=\frac{d}{da}\left(\frac{I}{I_0}\right)_0=\left(\frac{I}{I_0 a}\right)_{a=0}=\frac{RT_0}{\lambda M I_0},$$

and $$T_0=\frac{\lambda M I_0}{3R} \quad\text{.........................(15).}$$

On this theory if we write $f(a)$ for $\frac{\cosh a}{\sinh a}-\frac{1}{a}$ we have

$$I/I_0=f(MH_2/RT),$$

where f is the same for all substances. Since $H_2=\lambda I$ we have

$$\frac{I}{I_0}=f\left(\frac{\lambda M I}{RT}\right)=f\left(3\frac{T_0}{T}\frac{I}{I_0}\right),$$

so that $$I/I_0=\phi(T/T_0) \quad\text{.......................(16),}$$

where the function ϕ is the same for all substances. Thus if we express I, the intensity of permanent magnetization, in terms of the maximum possible intensity of magnetization I_0, and T the absolute temperature in terms of the absolute critical temperature T_0, we obtain a characteristic equation for the intensity of permanent magnetization which is identical for all ferromagnetic substances.

It is probable, as we shall see later, that the magnetic properties of the ferromagnetic metals are too much complicated by various secondary causes to afford a satisfactory test of the above theoretical conclusions. The properties of various crystalline ferromagnetic minerals, though complex enough of themselves, are in some respects simpler than those of the ferromagnetic metals and are better suited for carrying out a test of this character. The mineral magnetite has been found to be especially suitable for this purpose; but before discussing the experimental data which have been obtained with this substance we shall consider briefly the remarkable magnetic properties of another mineral, pyrrhotite, which seem likely to shed much light on the nature of ferromagnetic substances in general.

The Properties of Pyrrhotite.

The magnetic properties of this ferromagnetic mineral, which is a sulphide of iron having a composition very near to FeS, but with a slight excess of sulphur, have been investigated in detail by Weiss. The crystal has three mutually perpendicular planes of symmetry which may be indicated by the axes Ox, Oy, Oz. The crystals are much more easily magnetized parallel to one of these axes Ox than in any other direction, and furthermore the susceptibility parallel to Oy is much greater than that parallel to Oz. The plane xOy is called by Weiss the magnetic plane.

The magnetic phenomena exhibited by a uniform crystal when placed in a magnetic field parallel to Ox are characterized by remarkable simplicity. If the crystal shows no magnetic polarity to start with, the intensity of magnetization remains zero until H reaches a critical value $+H_c$, when the intensity of magnetization suddenly assumes its saturation value $+I_s$, which remains constant for all positive values of H and for all negative values greater than $-H_c$. As soon as $-H_c$ is reached the intensity of magnetization suddenly becomes $-I_s$, and retains that value until the field becomes $=$ or $> +H_c$. The I, H curve is thus a rectangle and the phenomenon is irreversible.

If the magnetizing field is inclined to Ox the phenomena are very different*. For different values of H the curves for I obtained by rotating the direction of H in the magnetic plane are shown in the accompanying figure. If H exceeds about 12,000 gauss the value of I is constant for all values of the field, but unless H is very great I is not in the same direction as H. This effect is well exhibited by the curve for $H = 4000$. The short lines represent the *direction* of the resultant magnetization for a field of intensity $H = 4000$ gauss inclined to the axis Ox at an angle given in degrees by the numbers alongside. For a variation in H of 5°

---

* P. Weiss, *Journal de Physique*, vol. IV. p. 486 (1905).

from the axis Oy the direction of the intensity of magnetization varies more than 45°. The component of the intensity of magnetization parallel to the field is a minimum when the angle θ between H and Ox is a maximum, for small fields. For larger fields the minimum occurs at intermediate values*.

The curves in Fig. 48 can be represented quite closely by a simple trigonometrical formula. If θ is the angle between the axis Ox and H, and ϕ that between Ox and I, then it appears that

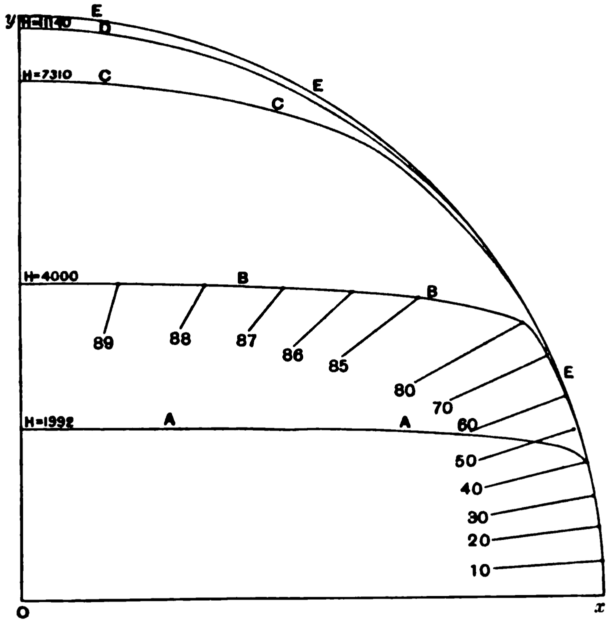


Fig. 48.

$$H \sin(\theta - \phi) - nI \sin\phi \cos\phi = 0 \quad \text{............(17)},$$

where n is a constant quantity. The phenomena in the plane xOz may be represented by a similar formula but with a different value of n.

A simple physical interpretation can be given to the foregoing equation. In addition to the field H the elementary magnets will be acted on by forces which depend upon the magnetization of the medium. These forces will certainly be partly magnetic, arising from the intermolecular magnetic field. They will not, at any rate necessarily, be entirely magnetic. They may arise as reactions, of elastic or other type, to the displacement of the elements of the material, from the normal equilibrium position, which is produced by the magnetic field. Since I measures a displacement of the elements from either a more normal or a less regular arrangement, the potential energy of the system which is due to these forces will be proportional to I^2 even though the forces are not actually magnetic. On account of the aeolotropy of the medium the coefficient of I^2 will be different along the different axes. Thus in general when H and I make angles θ and ϕ respectively with the axis of x, the x component of force per unit magnetic moment of an elementary magnet may be represented by $H \cos \theta + \lambda_1 I \cos \phi$ and the y component by $H \sin \theta + \lambda_2 I \sin \phi$, where λ_1 and λ_2 are constants depending on the structure of the material. Thus taking moments about the centre of mass of the elementary magnet

---

* P. Weiss, *loc. cit.* p. 487.

$$(H \cos \theta + \lambda_1 I \cos \phi) \sin \phi = (H \sin \theta + \lambda_2 I \sin \phi) \cos \phi,$$

or

$$H \sin (\theta - \phi) - (\lambda_1 - \lambda_2) I \sin \phi \cos \phi = 0.$$

This is the same as the former equation if $\lambda_1 - \lambda_2 = n$. For the plane xOz we have only to replace $\lambda_1 - \lambda_2$ by $\lambda_1 - \lambda_3$. It appears from the experiments that $(\lambda_1 - \lambda_2) I = 7300$ gauss and

$$(\lambda_1 - \lambda_3) I = 150{,}000 \text{ gauss},$$

whereas the maximum (saturation) value of I is about 47 gauss. The great difference between these numbers is somewhat surprising. If they were really magnetic forces one might expect them all to be of the same order of magnitude.

One of the most striking features of these phenomena is the very small field which can reverse the magnetization along the axis Ox compared with the fields which are required to produce any appreciable magnetization along the perpendicular axes. This is the more striking as the reversal of the magnetization would appear to involve the intermediate passage of the elementary magnets through the perpendicular orientation which is so difficult to produce, throughout the mass of the material, at any rate, by the application of an external field. There is not, however, any essential difficulty here. It seems clear that there are in general

We also saw on p. 382 that this theory led to a simple relation between the intensity of permanent magnetization and the temperature (equation (16), p. 383). This relation is the same for all substances, provided the intensity I_c is expressed in terms of the greatest possible value of I_c (the value at $T=0$) and the temperature T in terms of the temperature T_0 at which the ferromagnetism disappears: that is to say, the maximum value of I_c and the absolute temperature T_0 are to be taken as the units of intensity of magnetization and temperature respectively. This theorem has been tested experimentally by Weiss* in the case of magnetite. He finds that the theoretical curve is followed with great accuracy except in the immediate neighbourhood of $T=T_0$ and at very low temperatures. Even in these regions the deviations are not very large. The experiments extend from $-79°$ C. to the critical temperature $T_0=+587°$ C. When one considers the wide range of temperature covered, and the fact that there are no disposable constants in the formula, this agreement must be regarded as a remarkable confirmation of the theory.

There is one point which seems to call for further discussion at this stage. Equation (16) is derived from the equations

$$\frac{I}{I_0}=\frac{\cosh a}{\sinh a}-\frac{1}{a}, \quad a=\frac{MH}{RT} \text{ and } H=H_1+H_2$$

by putting the external field H_1 equal to zero and $H_2=\lambda I$, the derived equation (15) being used to eliminate the constants. Since I is a continuous function of T when $H_1=0$ it should also be a continuous function of H_1 and T when H_1 is allowed to vary. At first sight this appears to be contradicted by the experimental results. For, so far as the experiments have shown, there is no appreciable change in the value of $I=I_c$ as the external field H_1 is increased from zero to the highest values available in the laboratory. This would seem to be a fatal objection to the theory unless the values of the internal field λI were so great that the largest values of H_1 attainable were negligible in comparison. In that case the behaviour would be sensibly the same as for $H_1=0$ even in the highest fields which can be obtained. We shall see in the next paragraph that it is possible to deduce the values of the coefficients λ by an independent method, from experimental parameter is proportional to the field strength H. That this is at least approximately true is proved by the accompanying diagram, which represents the results of measurements by P. Curie on a specimen of iron in the neighbourhood of the critical temperature. The values of H are written alongside the corresponding curves.

---

* *Journal de Physique*, vol. VI. p. 665 (1907).

It is evident that equation (18) may be used to determine the values of the coefficient λ, since T, T_0, I and H_1 are given by the experiments. From measurements of this kind the following values of λ and H_2 have been deduced by Weiss:

Iron	$\lambda = 3{,}850$	$H_2 = 6{\cdot}56 \times 10^6$ E.M.U.
Nickel	$\lambda = 12{,}700$	$H_2 = 6{\cdot}35 \times 10^6$ „
Magnetite	$\lambda = 33{,}200$	$H_2 = 14{\cdot}3 \times 10^6$ „

The values of H_2 are enormous compared with the magnetic fields which can be obtained in the laboratory, so that the peculiar result that I_c does not appear to vary with the external field is accounted for satisfactorily.

Abrupt Changes in Magnetic Properties.

At high temperatures iron exhibits a number of more or less abrupt changes in its magnetic properties. Below 756° C. it exhibits the characteristic ferromagnetic properties usually associated with the metal. Between 820° C. and 920° C. it appears to be incapable of permanent magnetization, but it exhibits the rapid diminution of susceptibility with rising temperature which, as we have seen, should characterize ferromagnetic substances in the neighbourhood of the critical temperature. Between 920° C. and 1280° C. it behaves like a typical paramagnetic substance, the susceptibility varying as the inverse absolute temperature. Above 1280° C. it has similar properties, except that there is a sudden increase in the susceptibility at this temperature. This varying behaviour has been attributed to the existence of different modifications of iron within each of the limits of temperature specified. These modifications are denoted by Fe α, Fe β, Fe γ and Fe δ.

It has been pointed out by Weiss* that the constants which determine the magnetic behaviour of these different forms of iron exhibit simple numerical relations. These may be determined as follows:

---

* *Journal de Physique*, vol. VI. p. 685 (1907).

Fe α. In the neighbourhood of T_0 when I and a tend towards zero

$$I = I_0 \frac{a}{3} = \frac{I_0 MH}{3RT} = \frac{I_0 . NM . H}{3NRT},$$

where N is the number of magnetic molecules per unit volume of the material. But $NM = I_0$ and NRT is the pressure p which would be exerted by the substance if it were gaseous and occupied the same volume at the same temperature. Thus if δ is the density of the substance the constant

$$\chi = \frac{I}{H\delta} = \frac{I_0^2}{3p\delta}.$$

This formula is derived on the supposition that each kinetic molecule forms one magnetic molecule. If however n of the latter go to make up one of the former the right-hand side will have to be multiplied by n. Putting in known values of I_0 and δ, and the value of p derived from obvious data, one finds that

$$C = \chi T = 0{\cdot}00165n.$$

Fe β. We may attribute the quasi-ferromagnetic behaviour of this body to the fact that the external field H_1 is helped by the molecular field λI. Thus if $\chi\delta$ is the true susceptibility (δ = density) and $C\delta$ is the true value of Curie's constant

$$\frac{T}{C} = \frac{1}{\chi} = \delta \frac{H_1 + \lambda I}{I} = \frac{1}{\chi'} + \lambda\delta,$$

where $\chi'\delta$ is the measured value of the susceptibility. Comparing this with the equation

$$I(T - T_0) = \frac{H_1 T_0}{\lambda},$$

we see that, since H_1 is arbitrary,

$$C = T_0/\lambda\delta = \chi'(T - T_0).$$

Substituting the experimental value of $\chi' = 5{\cdot}13 \times 10^{-4}$ at 820° C. one finds

$$C = 0{\cdot}00164 \times 2.$$

Fe γ. The experimental values of the susceptibility of this paramagnetic body give the following values for C. At 940° C., $C = 0{\cdot}00172 \times 2$, and at 1280° C., $C = 0{\cdot}00182 \times 2$.

Fe δ. In a similar way the value of C for this paramagnetic body is found to be $0·00198 \times 3$ at 1280° C. and $0·00173 \times 3$ at 1336° C.

The interesting point is that all the values of C above are equal, within the limits of experimental error, either to twice or to three times the common factor 0·0017.

The occurrence of sudden changes in the magnetic qualities is by no means confined to iron. Similar features are presented by the other ferromagnetic substances which have been examined. In fact abrupt changes also occur even with diamagnetic substances. Thus Curie found that there was a large drop in the value of the diamagnetic constant of bismuth when fusion occurred. Moreover the diamagnetic constant of the molten bismuth was independent of the temperature, whereas this was not the case with solid bismuth. In the case of tin, which is sometimes diamagnetic and at others paramagnetic, according to the temperature, a number of transition points have been observed*. These facts support the view that a considerable part at least of the magnetic properties of bodies are determined by the occurrence of systems of considerable size rather than the atoms or sub-atomic structures†. It is of interest to form an estimate of the local strength of the molecular fields on the hypothesis that the apparent saturation of iron is due to the equilibrium between the internal fields and the kinetic energy of thermal agitation. Weiss (*loc. cit.* p. 688) estimates that for iron at ordinary temperatures $I/I_0 = 0·91$ about, whence, making use of Langevin's formula, $a = 11·3$. Thus

$$\frac{MH}{RT} = \frac{NM.H}{NRT} = \frac{NM.H}{p} = 11·3,$$

where p is the pressure exerted by the magnetic molecules, supposed gaseous, and filling the same space as the metal. Putting $NM = 2000$ gauss and $p = 2 \times 10^9$ dynes per sq. cm. one finds $H =$ about 11×10^6 lines per sq. cm. This rough calculation agrees as to order of magnitude with the more accurate estimate which was obtained on p. 390, and supports the conclusion to which we have already been driven that the behaviour of simple

---

* Du Bois and Honda, *Versl. Kon. Ak. van Wetensch. Amsterdam*, vol. XII. p. 596 (1910).

† Cf. Oxley, *Camb. Phil. Proc.* vol. XVI. p. 486 (1912).

For simplicity suppose that the magnetization arises entirely from the orbital motion of negative electrons whose charge is e. Let there be N of them per unit volume, and let A_z denote the average value of the projections of the areas of the orbits perpendicular to the axis of magnetization, divided by the times in which they are described. Then I_z, the magnetic moment per unit volume, is given by $I_z = NeA_z$. The resultant magnetization is taken to be parallel to the axis of z.

We shall now calculate the moment of momentum of the revolving electrons about the z axis. Consider any approximately circular orbit, the coordinates of whose centre are given by x_0, y_0, z_0. Let the coordinates of the revolving electron referred to this centre at any instant be ξ, η, ζ. The moment of momentum of the electron about the z axis is then

$$m\left\{(x_0+\xi)\frac{\partial\eta}{\partial t}-(y_0+\eta)\frac{\partial\xi}{\partial t}\right\}=m\left(x_0\frac{\partial\eta}{\partial t}-y_0\frac{\partial\xi}{\partial t}+\xi\frac{\partial\eta}{\partial t}-\eta\frac{\partial\xi}{\partial t}\right).$$

Averaging this over a complete revolution the mean values are

$$\overline{x_0\frac{\partial\eta}{\partial t}}=0=\overline{y_0\frac{\partial\xi}{\partial t}},$$

and $$\overline{\xi\frac{\partial\eta}{\partial t}}=-\overline{\eta\frac{\partial\xi}{\partial t}}=\frac{\text{area of projected orbit}}{t}=A_z.$$

Hence the average moment of momentum about any axis is independent of the position of that axis so long as its direction is the same. It is equal to $2mA_z$. The moment of momentum U_z per unit volume is

$$U_z = 2NmA_z = 2\frac{m}{e}I_z \text{ ....} (19).$$

It is thus equal to the intensity of magnetization multiplied by $2m/e$.

By the principle of the conservation of momentum the moment of momentum thus created must be balanced by an equal moment about the same axis. This reaction might conceivably occur either (1) on the electromagnetic system producing the exciting field, or (2) on the atoms of the magnetic material. In the former case the effect should depend simply on I_z and not involve e/m. In the latter case the effects observed depend on the looseness of the atoms. If they were free to rotate without affecting the neighbouring atoms, the moment of momentum of the orbits might be compensated by local rotations of the non-magnetic matter. In that case the magnetic material would not have any tendency to turn as a whole. On the other hand the fact that the magnetic properties of the material are very susceptible to changes of temperature makes it very unlikely that the connection between the magnetic atoms and the neighbouring matter is a loose one. It seems therefore most probable that the moment of momentum created in this way will be compensated by a motion of the magnetic material as a whole. Experiments which have been made to detect this effect have not led to a decision as to whether it exists or not.

Specific Heats of Ferromagnetic Substances.

If a substance is magnetized to an intensity I in a field of strength H, and I is proportional to H, the energy of the system is changed by an amount $\frac{1}{2}HI$. On account of the very large magnitude of the internal fields this energy is comparable with the whole thermal energy in the case of the ferromagnetic substances. This is true even when the average intensity of magnetization is zero. As the energy of the molecular fields is a function of the temperature, a very considerable part of the specific heat may arise in this way. The molecular magnetization diminishes with rising temperature so that additional energy must be supplied in order to overcome the attraction of the elementary magnets. The effect thus involves an increase in the specific heat of the substance. Since $H_2 = \lambda I$ this part of the specific heat is $\frac{\lambda I}{J}\frac{\partial I}{\partial T}$, where J is the mechanical equivalent of heat. Since I disappears suddenly at the critical temperature so will this additional specific heat. Weiss and Beck* have shown that the part contributed by the internal magnetic energy in this way accounts quantitatively for the anomalous specific heats of ferromagnetic substances.

---

* *Journal de Physique*, Ser. 4, vol. VII. p. 249 (1908).

reversible change occurring in an isolated system leaves the value of S unaltered, whereas any irreversible change increases S. Thus the final state is that for which S is a maximum. It may also be shown that the change in S produced by any reversible action depends only on the initial and final states of the system and not at all on the way in which the change has been effected, provided it is reversible throughout. Thus S is a perfect differential with respect to the independent variables in terms of which the state of the system is described.

The final state is only steady as regards quantities like pressure, temperature and so on, which are usually taken to be sufficient to describe the behaviour of matter in bulk. If we made use of instruments fine enough to determine the motions of the individual molecules there is every reason to believe that the system would be found to be the seat of very lively, never ceasing changes. The final steady state is therefore one of statistical equilibrium merely. The actual spatial distribution of the individual molecules and the distribution of the momentum and velocity which they possess are both constantly changing. On the average, however, and actually if the system contains an indefinitely large number of molecules, the distribution of the molecules in space and of momentum and velocity among the molecules is definite. We shall now attempt to discover what this law of distribution is.

The number of ways in which a given amount of energy may be distributed among an indefinitely large number of molecules is clearly infinite to a very high order. Some of these are much more probable than others and there will be one distribution which is the most probable. Jeans* has shown that in the statistically steady state which is independent of the time the most probable distribution is infinitely probable compared with the others; so that if we can find the most probable distribution we shall have obtained the actual distribution for all practical purposes.

Boltzmann pointed out the intimate connection between the probability of a given state of a system and the entropy of the system. In the final state of an isolated system we have seen that both the entropy and the probability of the system have attained a maximum value. In the intermediate stages both quantities are moving towards the maximum. Let us now introduce the hypothesis that the entropy of a system is a function of its probability only. This course is evidently permissible, since we have not yet defined the probability of a system precisely. We shall have to take care, when we come to do this, that our definition does not conflict with the hypothesis. Let us see what conclusions we may arrive at from the general conceptions of entropy and probability with the aid of our hypothesis.

---

* *Dynamical Theory of Gases*, Chap. III.

Consider two entirely separate material systems, let us say two stars so far apart that the interaction of their radiations may be disregarded. Then we have, by hypothesis, $S_1 = f(w_1)$ and $S_2 = f(w_2)$, where S_1 and S_2 are the entropies and w_1 and w_2 the probabilities of the systems separately. If S and w are respectively the entropy and probability of the two systems considered together, we have

$$f(w) = S = S_1 + S_2 = f(w_1) + f(w_2).$$

But $$w = w_1 w_2,$$

hence $$f(w_1 w_2) = f(w_1) + f(w_2),$$

so that $$S = k \log w \quad \text{......................} (2),$$

where k is a universal constant. Thus the difference between entropy and probability is only that one combines by addition and the other by multiplication.

Of the total number of molecules under consideration let us suppose that in the steady state the fraction

$$f(x, y, z, p, q, r)\, dx dy dz dp dq dr \quad \text{...........} (3)$$

have their centres between x and $x + dx$, y and $y + dy$, and z and $z + dz$, and their momenta between the components p and $p + dp$, q and $q + dq$, and r and $r + dr$. We shall suppose that the six variables x, y, z, p, q and r completely describe the state of the particles. This is equivalent to treating them as massive points subject to the action of forces, and although not general enough for many problems in dealing with gases, is sufficiently so for the electrical problems in which the particles under consideration are electrons. Another interpretation which may be given to $f(\quad)$ is that it is the probability that a particle selected at random should have its six coordinates within the assigned limits. We shall suppose that f is a continuous function which can be differentiated, the particles may be given the same position and momentum. This is evidently equal to one, which is also the value given by (4), since, for all the elements $d\sigma$ except the one in which the particles lie, we have $(nf\,d\sigma)! = 0! = 1$, and for the remaining element

$$(nf\,d\sigma)! = n!.$$

Since the size of the elements $d\sigma$ is arbitrary the expression for the probability above will only be the value in terms of some arbitrary standard. It is not necessary, however, for us to determine the value of the standard, since it is possible to arrive at results which are independent of $d\sigma$ without doing so.

By combining (4) with (2) we have

$$S = k \log n! - k \Sigma \log (nf\,d\sigma)! \quad \ldots\ldots\ldots\ldots\ldots(5).$$

Now we can always make $nf\,d\sigma$ as large as we please by taking n big enough. When N is any very large number we have Stirling's well-known approximation

$$\left.\begin{aligned} N! &= \left(\frac{N}{e}\right)^N \sqrt{2\pi N} \\ \text{or} \qquad \log N! &= N(\log N - 1) + \tfrac{1}{2}\log 2\pi N \\ &= N(\log N - 1) \end{aligned}\right\} \ldots\ldots\ldots\ldots(6)$$

with sufficient approximation, since we may neglect $\log N$ compared with N when N is large. Thus (5) may be written

$$\begin{aligned} S &= k \log n! - k \Sigma\, nf\,d\sigma\,(\log nf\,d\sigma - 1) \\ &= k \log n! - k \Sigma\, nf\,d\sigma\,(\log nf + \log d\sigma - 1) \\ &= k[\log n! - n(\log d\sigma - 1)] - k \Sigma\, nf\,d\sigma \,.\log nf, \end{aligned}$$

since all the $d\sigma$'s are equal and $\Sigma\, nf\,d\sigma = n$. Since n is constant we obtain

$$S = \text{const.} - k \int nf \log nf\,d\sigma \quad \ldots\ldots\ldots\ldots\ldots(7).$$

If the particles under consideration make up the whole of the system then (7) will be the complete expression for the entropy. Without any more elaborate analysis we may add to S a part which is independent of the particles under consideration and therefore independent of f. With this understanding we may put

$$S = S_0 - k \int nf \log nf\,d\sigma \quad \ldots\ldots\ldots\ldots\ldots\ldots(8)$$

On the other hand we obtain from (26)

$$\left(\frac{\partial S}{\partial L}\right)_V = \frac{3}{2}\frac{kn}{L} \text{ and } \left(\frac{\partial S}{\partial V}\right)_L = \frac{kn}{V},$$

whence, by comparison with (27),

$$pV = knT$$

and by comparing this with the equation of a perfect gas we have

$$k = R = (Tk_0)^{-1}.$$

The numerical value of R is known to the same degree of precision as that to which we know the value of the mass of an atom. With the accuracy which has recently been reached in this branch of physical measurement the different methods of determining R are consistent to within about 1 °/₀. One of the best values, deduced by Planck (see Chap. XV, p. 356) from the constants in the complete radiation formula, is

$$R = k = 1{\cdot}346 \times 10^{-16} \frac{\text{gm. cm.}^2}{\text{sec.}^2 \text{ deg.}}.$$

The Theory of Metallic Conduction.

The view, that the transportation of electricity in metallic conductors, like the corresponding phenomenon exhibited by electrolytes, is due to the motion of minute charged particles, was a definite feature of the old Weberian formulation of electrodynamics. The modern development is due to the labours of Riecke, Drude, J. J. Thomson and others. According to the most prevalent form of the theory, and the one which we shall adopt, the atoms of the metal are regarded as continually dissociating into a negative electron and a positively charged residue. It follows from the kinetic theory considerations which we have just discussed that the electrons will be moving in all directions in the interior of the metal with very high speeds. They will in fact possess the same distribution of velocity as would an uncharged molecule of equal mass, and their mean kinetic energy will be proportional to the absolute temperature. We should rather expect that the atoms and positively charged residues would have an equal distribution of kinetic energy, but it is possible that they are jammed in some way which prevents this coming about. Whether the atoms and positive residues are able to move or not, they are to be regarded as only oscillating about fixed positions and not travelling from point to point of the material.

We now suppose that the effect of applying an electromotive intensity is to superpose on the haphazard motion of the electrons, which arises from thermal agitation, a velocity of drift which, since they are negatively charged, is in the opposite direction to the applied force. This motion constitutes the electric current, which is thus carried entirely by the negative electrons. *A priori*, one might be tempted to suppose that the positive residues would also drift along under the influence of the electric field. If any such effect exists it must be extremely small. For if it were appreciable we should expect an electric current to transport atoms of one metal into the other across the junction between two metals. Careful experiments have been made to detect such effects, but they have always led to negative results. This objection would be removed if we supposed the positive particles to be of the same nature in different materials. There are, however, two serious objections to such a view. In the first place there is no other evidence of the liberation of such particles from atoms under conditions which can be considered at all analogous to those which hold inside conductors. In the second place the hypothesis appears to be incapable of removing such difficulties as are presented by the simpler theory.

The strongest arguments in favour of the view which asserts that the currents in metals are carried by negative electrons are as follows:

(1) Conductors when heated or illuminated are found to emit electrons into the surrounding space. The mass of the electrons thus emitted from a hot wire may, in favourable cases, be large compared with the number which, on any reasonable hypothesis, may be expected to be present in the wire at any instant, showing that electrons are continuously flowing into the wire from other parts of the system*.

(2) The typical good conductors at low temperatures are all metals: i.e. electropositive elements which are known from other phenomena to liberate electrons from their atoms readily.

---

* **Cf. O. W. Richardson, *Phil. Mag.* vol. xxvi. p. 345 (1913).**

Material	Ratio: Thermal Conductivity to Electrical Conductivity	Temperature Coefficient of this Ratio, per cent.
Copper, commercial	$6{\cdot}76 \times 10^{10}$ at 18° C.	—
Copper (1) pure	$6{\cdot}65 \times 10^{10}$ at 18° C.	·39
Copper (2) pure	$6{\cdot}71 \times 10^{10}$ at 18° C.	·39
Silver, pure	$6{\cdot}86 \times 10^{10}$ at 18° C.	·37
Gold (1)	$7{\cdot}27 \times 10^{10}$ at 18° C.	·36
Gold (2) pure	$7{\cdot}09 \times 10^{10}$ at 18° C.	·37
Nickel	$6{\cdot}99 \times 10^{10}$ at 18° C.	·39
Zinc (1)	$7{\cdot}05 \times 10^{10}$ at 18° C.	·38
Zinc (2) pure	$6{\cdot}72 \times 10^{10}$ at 18° C.	·38
Cadmium, pure	$7{\cdot}06 \times 10^{10}$ at 18° C.	·37
Lead, pure	$7{\cdot}15 \times 10^{10}$ at 18° C.	·40
Tin, pure	$7{\cdot}35 \times 10^{10}$ at 18° C.	·34
Aluminium	$6{\cdot}36 \times 10^{10}$ at 18° C.	·43
Platinum (1)	$7{\cdot}76 \times 10^{10}$ at 18° C.	—
Platinum (2) pure	$7{\cdot}53 \times 10^{10}$ at 18° C.	·46
Palladium	$7{\cdot}54 \times 10^{10}$ at 18° C.	·46
Iron (1)	$8{\cdot}02 \times 10^{10}$ at 18° C.	·43
Iron (2)	$8{\cdot}38 \times 10^{10}$ at 18° C.	·44
Steel	$9{\cdot}03 \times 10^{10}$ at 18° C.	·35
Bismuth	$9{\cdot}64 \times 10^{10}$ at 18° C.	·15
Constantan (60 Cu, 40 Ni)	$11{\cdot}06 \times 10^{10}$ at 18° C.	·23
Manganin (84 Cu, 4 Ni, 12 Mn)	$9{\cdot}14 \times 10^{10}$ at 18° C.	·27

ratio is very close in all cases to the theoretical value. The deviations are greater for the poorer conductors and in almost all cases are in the direction of values greater than the theoretical. This is what would happen if the part of the thermal conductivity which does not depend on the motion of the electrons were to become appreciable. Thus the deviations lie in the direction in which they would be expected to occur. The behaviour of alloys is exceptional, but so are most of their electrical properties. In fact Lord Rayleigh* has pointed out that the electrical resistance of alloys may be expected to be unduly high on account of the existence of a "false resistance" arising from the Peltier heating effect at the junction of parts of the material of varying composition. It is also desirable to mention that Lees† has found that the divergence of the values of $k/\sigma T$ for different substances is greater at the temperature of liquid air than at ordinary temperatures. On the whole, however, the concordance of the values of this quantity for so many metals over so wide a range of temperatures is more striking than the differences, when one considers the number of factors which might enter.

The percentage temperature coefficient required by the theory of the colliding particles. Thus, so far as the set of assumptions we are dealing with is concerned, a complete theory would have to take into account the occurrence of collisions between electrons, or something which produces the same effect, in order to account for the existence of Maxwell's distribution under normal conditions. The mathematical difficulties of a more complete theory, unfortunately, appear insuperable. Our method of attack will be to try to find a law of distribution of velocity, slightly different from the normal one, which will make the distribution at every point a steady one when electric or other forces act on the electrons in the metal.

---

* *Nature*, vol. LIV. p. 154; *Scientific Papers*, vol. IV. p. 232.

† *Phil. Trans.* A. vol. CCVIII. p. 381 (1908).

Consider the causes which tend to change the number of electrons, at any point, which have assigned components u, v, w of velocity. These are two in number, viz. (1) the free motion of the electrons from one part of the conductor to another, and (2) collisions. Let us imagine the distribution of velocity among the electrons *at any point* to be represented by a three-dimensional velocity diagram. The diagram is drawn so that the resultant velocity of each electron at the point is represented by a radius from the origin. The density of the points which are the ends of such lines and which lie within any element of volume $d\sigma = du\,dv\,dw$ of this diagram will represent the number of electrons at the given point which have velocity components between u and $u + du$, v and $v + dv$, and w and $w + dw$. In this discussion we mean, of course, by the expression the number "at any point" x, y, z the number in an infinitesimal element of volume $d\tau = dx\,dy\,dz$ which contains the point x, y, z. Let us denote the number of electrons which have velocity components u, v, w at the point x, y, z at the instant t by $f\,(u, v, w, x, y, z, t)\,d\sigma\,d\tau$. If the electric intensity in the metal is X it will give rise to an acceleration of each electron equal to Xe/m. X is supposed to be parallel to the x axis. If there were no collisions these electrons would be found in a different element $d\sigma'\,d\tau'$ of the velocity and space diagram at a later instant $t + dt$. Owing to the motion the new velocity coordinates would evidently be $u + X\dfrac{e}{m}\,dt$, v, w, and the space coordinates $x + u\,dt$, $y + v\,dt$ and $z + w\,dt$ instead of u, v, w, x, y, z respectively. Corresponding points would be displaced equally in each diagram so that $d\sigma' = d\sigma$ and $d\tau' = d\tau$. If there were no collisions we should deflections take place the influence of one of the centres is very great compared with that of all the others. As the term collision is often understood to imply the occurrence of geometrical contact, like that between hard elastic spheres, we shall replace it by the more general term encounter, throughout the rest of this discussion. The sequence of changes which characterize an encounter is then to be represented mathematically as follows. As the previous history of the electron cannot affect the results of the calculation we can suppose it to have been moving for an indefinite time with the uniform velocity V in a straight line. As the encounter begins to occur the linear path becomes curved, but the magnitude of V is unchanged. As we are neglecting all encounters in which more than one centre of force plays an important part, the orbit of the electron will lie entirely in one plane, that which contains the direction of its original motion and the perpendicular from the centre of force upon it. If the orbit is an open one it will end by becoming asymptotic to a straight line inclined at an angle, which we shall call 2θ, to the direction of the original straight path. θ is evidently the deflection up to the apse, about which the orbit is symmetrical. If the perpendicular distance from the deflecting centre to the original straight path is b, the number of collisions made in unit time by a single particle whose speed is V, such that the distance b lies between b and $b+db$ and the plane of the orbit lies in an azimuth between ψ and $\psi+d\psi$, where ψ is measured from a fixed plane passing through the direction of V, which is taken as the axis about which ψ is measured, is $nrb\,db\,d\psi$, where n is the number of the deflecting centres in unit volume of the substance. This follows since the expression above is equal to the number of centres in the region between two circular cylinders of radii b and $b+db$, whose height is equal to the velocity V of the moving particle, which is cut off by the two planes ψ and $\psi+d\psi$ which pass through V and determine the plane of the orbit. Since there are $f(u, v, w)\,d\sigma$ particles with velocity components about u, v, w in unit volume, it follows that the number of particles which leave the group u, v, w, x, y, z in unit time owing to deflections through angles which lie between 2θ and $2(\theta+d\theta)$ in an azimuth between ψ and $\psi+d\psi$ is

$$nVf(u, v, w, x, y, z, t)\,b\,db\,d\psi\,d\sigma\,d\tau\,dt \,\ldots\ldots\ldots (32).$$

T^{-1} might arise from an appropriate variation of one or more of the quantities N, n and K with T. At present there is no means of finding the connection of K, N and n separately with T.

Recent experiments by Kammerlingh Onnes and his collaborators have shown that the resistance of the pure metals becomes exceedingly small at very low temperatures. This result cannot be said to be definitely indicated by the formulae which we have just developed. It is, however, not necessarily in conflict with them. For example we might identify the centres, with which the electrons collide, with the vibrators which take part in thermal phenomena. The most obvious physical interpretation of these vibrators is to regard them as electrical doublets. The doublets might arise, for example, by the vibration or steady motion of electrons about an equilibrium position. They would exert forces on the moving electrons which would vary inversely as the cube of the mutual distance, a result which, as we have seen, is required by several considerations. It follows from Einstein's theory of specific heats that the energy of the vibrators approaches zero exponentially as the absolute zero of temperature is approached. Now the average moment of a doublet constituted in this way is proportional to the square root of its energy: so that the moment or strength of the centres will also approach zero exponentially as the temperature is reduced. Since there is no reason for supposing that the number of free electrons approaches zero at so fast a rate as this, the conductivity would become infinite at very low temperatures.

Such a view is in satisfactory agreement with the experimental measurements. By assuming that K times the product on the right-hand side of (47) varies as $T^{-\frac{1}{2}}$, which gives the right variation of σ with T at high temperatures, and that the variation of K with T depends upon the contained heat energy, as deduced from the specific heat measurements, in the manner just indicated, I find that the calculated temperature at which the electrical resistance disappears agrees quite accurately with that given by Kammerlingh Onnes in the case of mercury and is not far from the values given for lead and gold. The way in which the calculated electrical resistance varies with the temperature is also very similar to that found experimentally at these low temperatures. The same results would not follow from Planck's later inclusion of a single wire of the material B, the temperature of which varies from one extreme to the other and back again. As inequality of temperature in a closed circuit of a single uniform material does not give rise to any thermoelectromotive force, this cannot cause any difference from the effect given by AC simply.

In addition to the thermoelectromotive force caused by a difference of temperature there are the converse reversible heating and cooling effects which are produced by the flow of an electric current. We have thus to consider the Peltier effect, which is the heat liberated when an electric current flows across the junction between two different metals, and the Thomson effect, which is the heat developed reversibly when an electric current flows along an unequally heated bar. These are measured by the respective coefficients P and σ. Both of these refer to the amount of heat liberated by the passage of unit quantity of electricity. In specifying σ the electricity has to flow against one degree difference of temperature, the directions of the electric and thermal gradients being coincident. σ is positive when positive current flowing in the direction of increasing temperature causes an absorption of heat.

The application of the conservation of energy to a thermoelectric circuit gives

$$Ei = Ri^2 + i\Sigma P + i\int \sigma\, dT$$

if R is the resistance of the circuit. When i is made very small the Joulian development of heat Ri^2 vanishes in comparison with the other terms, so that the reversible quantities satisfy the equation

$$E = \Sigma P + \int \sigma\, dT \quad\text{..........................(49).}$$

Similarly by applying the second law of thermodynamics in the form $\int \frac{dQ}{T} = 0$ to the reversible heat production, we have

$$\Sigma \frac{P}{T} + \int \frac{\sigma}{T}\, dT = 0 \quad\text{..........................(50).}$$

It follows from these equations that

$$\frac{\partial E}{\partial T} = \frac{P}{T} \quad \text{and} \quad \sigma_B - \sigma_A = T\frac{\partial}{\partial T}\left(\frac{P}{T}\right) \quad\text{..........(51).}$$

$$W = \frac{2\pi}{3} m \left\{\left(2hAX - \frac{dA}{dx}\right) J_6 + A \frac{dh}{dx} J_8\right\} \text{ ....(54)}$$

$$= -\frac{RT}{e^2} \sigma_0 \mu \left\{mX - RT \frac{d \log A}{dx} + \mu' R \frac{dT}{dx}\right\} \text{ ....(55)}$$

$$= -\frac{RT}{e} \mu i - k \frac{dT}{dx} \text{ ....(56),}$$

where

$$J_n = \int_0^\infty V^n \psi e^{-hV^2} dV,$$

$$\sigma_0 = \text{specific electrical conductivity} = \frac{4\pi}{3} \frac{e^2}{RT} A J_4,$$

$$\mu = -\tfrac{1}{2} \frac{m}{RT} J_6/J_4, \quad \mu' = -\frac{m}{2RT} J_8/J_6,$$

$$k = \text{thermal conductivity} = \frac{\pi}{3} \cdot \frac{A}{R} \frac{m^2}{T^2} \left\{J_8 - \frac{J_6^2}{J_4}\right\},$$

and $$k/\sigma_0 = \frac{m^2}{e^2} \frac{1}{T} \frac{1}{4} \left\{\frac{J_8}{J_4} - \left(\frac{J_6}{J_4}\right)^2\right\} = \frac{R^2}{e^2} T\mu (\mu' - \mu).$$

If the centres are far enough apart and act on one another with forces which vary as the inverse sth power of the distance, we have, as before,

$$\mu = -\left(\frac{2}{s-1} + 2\right) \quad \text{and} \quad \mu' = -\left(\frac{2}{s-1} + 3\right).$$

In the equations above, X is the mechanical force on an electron. If X_0 is the electric intensity, then $eX_0 = mX$.

The thermoelectromotive force E round any circuit is the value of $\int X_0 dx$ which is required to reduce the current i to zero, when there are no batteries in the circuit. Thus from (53)

$$E = \int X_0 dx \quad \text{when} \quad i = 0$$

$$= \frac{R}{e} \int \left\{T \frac{d \log A}{dx} - \mu \frac{dT}{dx}\right\} dx$$

$$= \frac{R}{e} \left\{ | T \log A | - \int (\log A + \mu) \, dT \right\} \text{ ....(57),}$$

after integrating by parts. If the integral is taken round a closed circuit of two metals with varying temperatures, the integrated

(53) that when there is no difference of temperature $\left(\frac{\partial T}{\partial x}=0\right)$, as i approaches zero X_0 approaches $\frac{RT}{e}\frac{\partial \log A}{\partial x}$; so that the work done by the electric force at the boundary is $RT \log \frac{A_1}{A_2}$. The difference between this and eP_{12} is due to the fact that, when the law of force is different in different metals, the amount of kinetic energy transported by the electrons which carry a given current is different. As the ratio of the thermal and electric conductivities indicates that the law of force is very nearly the same for the pure metals which are good conductors, it is probable that $\mu_1 - \mu_2$ is not large. It may, however, be of the same order of magnitude as the observed Peltier and Thomson effects.

The values given by (58), (59), (61) and (62) satisfy the equations (49)—(51) of Lord Kelvin's thermodynamic theory. A different method of deducing some of the thermoelectric formulae will be considered in the next chapter.

Conductivity for Periodic Forces.

The behaviour of the electrons in metals under alternating forces has been considered by various writers. The following investigation, due to H. A. Wilson*, is an elaboration of a method originally given by Jeans†. The electrons instantaneously present in any given volume may be divided into groups characterized by a particular velocity of agitation V. The number of collisions of the electrons with one another is regarded as small compared with the number of collisions between electrons and atoms. Owing to the relatively small mass of the electrons the latter class of collisions will have very little effect in changing the magnitude of the velocity V. Thus the electrons in any given group will remain in that group throughout a large number of collisions. Let dN denote the number of electrons in the group characterized by the value V and let u denote the average component of velocity of this group in the direction of the electric intensity X. Consider the rate at which the momentum of the group is changing. The group is gaining momentum from the applied field at a rate which chemical action. The electrons which are liberated are then to be regarded as one of the products, either intermediate or final, of the chemical action, and the extent to which they are formed will be determined by the laws which govern the formation of other chemical products. Now, the law which governs the rate of chemical actions in general is not essentially different from (I), and something very like the other results enumerated on p. 441 would also follow in this case, so that such considerations will not enable us to distinguish between the two phenomena. The only satisfactory criterion is whether the emission of electrons is accompanied, *pari passu*, by chemical combination or decomposition of the more generally recognized type. There is no evidence that this is universally the case.

---

* *Phil. Mag.* VI. vol. xx. p. 835 (1910).

† *Ibid.* vol. XVII. p. 773 (1909).

We shall now return to the consideration of the theory of the purely thermal emission. This phenomenon may be attributed to the increased kinetic energy of the electrons at high temperatures enabling them to overcome the forces which tend to retain them in the conductor. The rate of variation of the emission with the temperature may be calculated in a variety of ways. It will be conducive to clearness if we sacrifice generality to a slight extent and make a definite hypothesis about the structure of the interior of a metallic conductor: although many of our results will be more general than the hypothesis we are making. We have seen that the electrons in a good conductor behave as though they were acted upon by fixed centres of force varying as the inverse third power of the mutual distance. The potential energy of an electron must therefore be continually varying from point to point of its path. We shall suppose that the potential energy of an electron inside a metal is a function of its position only. In other words, we are supposing that the interior of a metal may be sufficiently described by mapping it out by a series of fixed surfaces, the level surfaces of W, the potential energy of an electron. These surfaces are supposed to be definite and characteristic for each conductor. The fact that the electrons are in motion prevents this from being a complete representation of the state of affairs, but it will be fairly certain to give a satisfactory account of the more important features of the phenomena.

The state of the electrons at any point of such a system as we are considering is defined by the average number ν of them per unit volume at that point at any instant, by the potential energy W of the electrons at that point, and by the temperature T, which is proportional to the average kinetic energy of the electrons. At points just outside any conductor W takes the constant value W_0. In the state of equilibrium the values of W_0 are different and characteristic for different conductors. This definition of W_0 is not sufficiently exact. At points very close to the conductor an electron is attracted towards it by its mirror image in the conductor. Thus W increases for some little distance away from the conductor and the points at which it approaches a sensibly constant value are not immediately outside the bounding surface. On the other hand, at considerable distances from the surface W will be affected by the potential of other bodies in the neighbourhood. The true value of W_0 is the value of W at points at a considerable distance from the surface, either in a cavity inside the conductor or outside if there are no other bodies in the neighbourhood.

It follows, as a result of our investigation of the kinetic theory of matter (p. 404), that at any point in a system in equilibrium at temperature T,

$$dn = \nu\, d\tau = K e^{-W/RT}\, d\tau \quad \ldots\ldots\ldots\ldots\ldots\ldots(1),$$

where dn is the number of electrons, which participate in thermal phenomena, in the element of volume $d\tau$, and K is constant throughout the system, being a function of T only.

We shall now consider the formulae which determine the equilibrium between the external free electrons and the internal electrons which can become free. Confining our attention to a single conductor, consider first the special case in which there are a finite number p of finite internal regions each characterized by constant values ν_1, W_1, τ_1, ... ν_p, W_p, τ_p. Then if ν_0, W_0, τ_0 are the values of the corresponding variables just outside the conductor, we have

$$\frac{\nu_0}{e^{-W_0/RT}} = K = \frac{\nu_1}{e^{-W_1/RT}} = \ldots = \frac{\nu_p}{e^{-W_p/RT}} = \frac{n}{\Sigma_\tau e^{-W/RT}} \ldots\ldots(2),$$

where n is the total number of electrons which can become free that are present in the system. In general it will be possible to regard W as having a constant value only over a region of

$$\frac{\partial J}{\partial v} = \underset{\delta v = 0}{\mathrm{Lt}} \frac{1}{\delta v} \{ J_{\tau+v+\delta v} - J_{\tau+v} \}$$

$$= \frac{e^{-W_0/RT}}{\int_{\tau+v} e^{-W/RT}\, d\tau} \left\{ W_0 - \frac{\int_{\tau+v} W e^{-W/RT}\, d\tau}{\int_{\tau+v} e^{-W/RT}\, d\tau} \right\} \ldots(10),$$

where $\int_{\tau+v}$ denotes that the integral is taken over the volume τ of the conductor plus the volume v of the external space.

Let n_v be the number of electrons in the external volume v and n_τ the number, of the kind contemplated, in the metal. Then in these problems n_v is always extremely small compared with n_τ. And since

$$\frac{n_\tau}{\int_\tau e^{-W/RT}\, d\tau} = K = \frac{n_v}{\int_v e^{-W/RT}\, d\tau}$$

by (3), we may replace

$$\int_{\tau+v} e^{-W/RT}\, d\tau \text{ in (10) by } \int_\tau e^{-W/RT}\, d\tau.$$

Also since

$$\frac{\int_v W e^{-W/RT}\, d\tau}{\int_\tau W e^{-W/RT}\, d\tau} = \frac{J_v}{J_\tau} \frac{n_v}{n_\tau} = \frac{J_v}{J_\tau} \frac{\frac{n_v}{v}}{\frac{n_\tau}{\tau}} \frac{v}{\tau} = \frac{J_v}{J_\tau} \frac{\nu_v}{\nu_\tau} \frac{v}{\tau},$$

and the following equation is true as regards order of magnitude,

$$\frac{\nu_v}{e^{-J_v/RT}} = K = \frac{\nu_\tau}{e^{-J_\tau/RT}},$$

the order of magnitude of

$$\frac{\int_v W e^{-W/RT}\, d\tau}{\int_\tau W e^{-W/RT}\, d\tau}$$

is that of $\frac{n_v}{n_\tau}$. Thus $\int_{\tau+v} W e^{-W/RT}\, d\tau$ may also be replaced by

The Relation with the Specific Heat of Electricity.

There is an intimate connection between ν_0, T, w and the specific heat of electricity σ within the conductor under consideration. Suppose that we have two conductors A and A' made of the same material, but maintained at the different absolute temperatures T and T'. They are to be of sufficiently great size and are connected by a thin conductor of the same material. The atmospheres of electrons about A and A' are enclosed and separated from each other by a suitable insulating boundary (see Fig. 50).

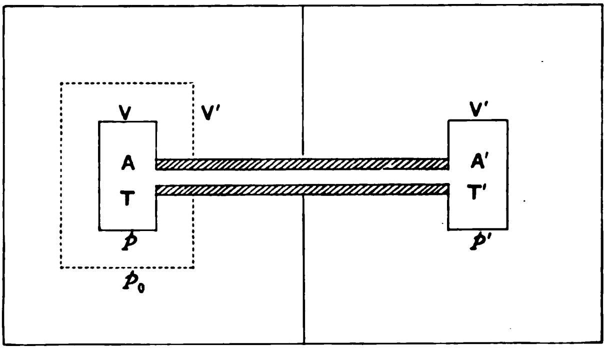


Fig. 50.

If contact difference of potential depends upon temperature, the potential V' at the surface of A' will not be the same as that at the surface of A. If the value at the latter surface is V, the electrons in the enclosure surrounding A will be at a different electric potential from those surrounding A'. Let eV' be $> eV$, and surround A with a screen which is permeable to the electrons and maintained at the potential V'. This may be imagined as a wire gauze of indefinite fineness which is connected to one end of a battery, the other end being connected to A. The electromotive force of the battery is $V' - V$. There is no work done by the battery, but even if there were it would not affect the argument. The nett effect of this arrangement is that the electrons in the enclosure outside the potential filter have the temperature T and the potential V'. Their pressure p_0 is different from the pressure p outside A, the relation being evidently

$$\log p_0 = \log p - \frac{e(V' - V)}{RT} \quad \text{...............(15)}.$$

Approximation to ν_0 as a function of T.

If the electrons are treated as point charges, they have three degrees of freedom and $\gamma = \frac{5}{3}$. If the collisions of the electrons occur with centres of force varying as a power of the mutual distance, it follows from the kinetic theory calculations in the last chapter that $\frac{d}{dT}\left(\frac{\lambda_1 - \lambda_0}{T}\right) = 0$, so that, under these conditions,

$$\sigma = \frac{1}{e}\left\{\frac{3}{2}R - \frac{\partial\phi}{\partial T}\right\} \quad \text{.....................(25).}$$

The value of σ is different for different metals and may be either positive or negative. In general it also varies with the temperature, often in a complicated manner. If the thermoelectric powers of different pairs of metals were accurately linear functions of the

Metal	Temperature ° C.	σ Erg/El. mag. C.G.S. unit	Observer
Mercury	+ 50	$6{\cdot}8 \times 10^2$	Schoute
	+100	$8{\cdot}6 \times 10^2$	,,
	+150	$10{\cdot}6 \times 10^2$	,,
Copper	−100	$0{\cdot}9 \times 10^2$	Berg
	0	$1{\cdot}6 \times 10^2$	,,
	+100	$2{\cdot}0 \times 10^2$	,,
	+300	$2{\cdot}1 \times 10^2$	Locher
	+500	$2{\cdot}6 \times 10^2$	,,
Silver	+100	$3{\cdot}46 \times 10^2$	,,
	+300	$4{\cdot}20 \times 10^2$	,,
	+500	$4{\cdot}95 \times 10^2$	,,
Platinum	− 50	$-9{\cdot}4 \times 10^2$	Berg
	0	$-9{\cdot}1 \times 10^2$	,,
	+ 50	$-9{\cdot}0 \times 10^2$	,,
	+100	$-9{\cdot}1 \times 10^2$	,,
Iron	0	$-4{\cdot}0 \times 10^2$	,,
	+100	$-12{\cdot}4 \times 10^2$	,,
	+100	$-13{\cdot}8 \times 10^2$	Locher
	+200	$-16{\cdot}8 \times 10^2$	,,
	+300	$-14{\cdot}2 \times 10^2$	,,
	+400	$-7{\cdot}5 \times 10^2$	,,
Constantan	0	$-23{\cdot}0 \times 10^2$	,,
	+200	$-19{\cdot}9 \times 10^2$	,,
	+400	$-13{\cdot}7 \times 10^2$	,,

If we apply a sufficiently great potential difference all the N electrons emitted by the conductor may be collected on a neighbouring electrode. The ensuing saturation current is, per unit area of the emitting body,

$$Ne = A_1 \left(\frac{R}{2\pi m}\right)^{\frac{1}{2}} T^2 \epsilon^{-\phi_0 J/RT} = A_2 T^2 \epsilon^{-\phi_0 J/RT} \quad \text{......(29)},$$

where A_2 and ϕ_0 are constants characteristic of the substance but independent of T. Had we not made the approximations above the index of T would be slightly different from 2 and we should have had exponential factors of the type $\epsilon^{a_p T^p}$, where a_p is an extremely small constant. The values of these factors cannot differ much from unity. We should therefore expect (29) to give a fair representation of the mode of variation with temperature of the electrons emitted by hot bodies. The number of substances which have been tested in this way is now very considerable and for all of them the emission has been found to be consistent with an equation of the form $Ne = AT^{\frac{1}{2}}\epsilon^{-b/T}$. As, however, the variation is almost all in the exponential term the results can be fitted as well by $Ne = AT^2\epsilon^{-b/T}$, by taking a slightly different value of the constant b in the exponential index.

We shall now consider the bearing of these effects on the nature of contact electromotive force.

Contact Difference of Potential.

Imagine an enclosure limited by an insulating boundary, maintained at temperature T. Suppose the enclosure to contain q material bodies, arranged in any manner, and that the whole system has come to a state of thermal equilibrium. The temperature will then be uniform and equal to T throughout the system. In general the surfaces of the bodies will assume different potentials $V_1, V_2, \dots V_q$. Clearly

$$e(V_m - V_p) = W_0^m - W_0^p \quad \text{.................(30)},$$

where W_0^m is the potential energy of an electron just outside the mth body. If we can show that $W_0^m - W_0^p$ are uniquely determined by the constitutive equations of the system it will follow that $V_m - V_p$ represent true intrinsic differences of potential. Let ν_0'

etc. represent the concentrations of the electrons at points just outside the different bodies. Since there is equilibrium between the electrons which are just outside each pair of bodies we evidently have $q-1$ equations of the type

$$\nu_0^m/\nu_0^p = \epsilon^{-\frac{W_0^m - W_0^p}{RT}}.$$

From (14) and (12) we see that there are q equations of the type $\nu_0^m = \phi_m(W_0^m J_m A_m T)$. From (7) there are also q equations of the type $J_m = \psi_m(T)$ and evidently there are q equations of the type $A_m = \text{constant}$. Altogether there are $4q-1$ equations between the $3q$ variables ν_0, J and A, the $q-1$ differences $W_0^m - W_0^p$ and T. Thus there are $4q-1$ equations and $4q$ variables, so that if one of the latter, say T, is given all the variables including the $q-1$ differences $W_0^m - W_0^p$ are determined. These equations involve neither the size, shape nor relative orientation of the bodies, so that the differences of potential $V_m - V_p$ are characteristic of the substances under consideration. They are clearly the same whether the bodies are in contact or not.

Since from (13) $w = W_0 - J$ it follows that

$$W_0^m - W_0^p = w_m - w_p + J_m - J_p.$$

We shall see below (p. 457) that $J_m - J_p$ is small compared with $w_m - w_p$; so that according to this view the contact difference of potential may be estimated from the internal latent heat of evaporation of the electrons. The experimental determinations of w for different substances are not yet sufficiently trustworthy to furnish a satisfactory test of this relation; although such indications as there are tend to show that w is smaller for electropositive than for electronegative elements.

The Peltier Effect.

We shall now proceed to obtain expressions for the various physical quantities which are grouped under thermoelectric phenomena. Let us first consider the Peltier coefficient P. To do this we need only to consider a reversible isothermal cycle involving two different conductors. The arrangement and operation is in fact much the same as that shown in Fig. 50 except that the conductors A and A' are of different materials and are at the same temperature T. Let the suffixes 1 and 2 be used to denote the various physical quantities which relate to the separate materials, the notation being otherwise as before. Let eV_2 be $> eV_1$. Surround the first conductor, as before, by a potential filter maintained at V_2. This will reduce the pressure of the electrons from p_1 to $p_1{}^0$ and will change their potential energy from eV_1 to eV_2 but will not affect their temperature T. The cycle commences by removing N electrons from the enclosure surrounding A_1 under the uniform pressure $p_1{}^0$. The work done is NRT and the heat absorbed is, by virtue of (13), $N[w_1 + e(V_2 - V_1) + RT]$. Next expand at T from $p_1{}^0$ to p_2. The work done and heat absorbed are each equal to $NRT \log (p_1{}^0/p_2)$. Next condense at p_2 and V_2 in the second body. The work done is $-NRT$ and the heat absorbed is $-N[w_2 + RT]$. The electrons are then allowed to flow along the connecting conductor to A_1, thus completing the reversible cycle. During this operation no work is done, but there is an absorption of heat NQe at the junction. The total amount of work in this isothermal reversible cycle is $NRT \log (p_1{}^0/p_2)$. Since the cycle is isothermal this must vanish, so that $p_1{}^0 = p_2$. The total absorption of heat is $N[w_1 + e(V_2 - V_1) + RT + RT \log(p_1{}^0/p_2) - w_2 - RT + Qe]$, and since this also must vanish we have

$$Qe = w_2 - w_1 - e(V_2 - V_1) \quad \text{..............(31).}$$

If Q is to be the same thing as the coefficient P which is measured in experiments on the Peltier effect, it is necessary that the cycle should be operated under the same conditions as when a current flows continuously under an impressed electromotive force. Hence, as in dealing with the Thomson effect, w_2 and w_1 will be the values of the latent heats which correspond to a steady flow and not to equilibrium conditions. Denoting the former by w and the latter by ϕ as before, we therefore have the equations

$$\begin{aligned} eP &= w_2 - w_1 - e(V_2 - V_1) \quad \text{.......................(32),} \\ &= \phi_2 - \phi_1 - (\lambda_2 - \lambda_1) + e(V_2 - V_1) \quad \text{.........(33),} \\ &= J_1 - J_2 - (\lambda_2 - \lambda_1) \quad \text{.........................(34),} \end{aligned}$$

since $\qquad \phi \equiv W_0 - J.$

A series of experimental values of P as well as the values of the thermoelectric power of the same pairs of metals at the same temperatures are given in the following table, taken from Baedeker[*].

Materials	Temperature C.	Peltier coefficient in		Thermoelectric[†] power. $1 = 10^{-6}$ volt per degree C.
		Millicalories per coulomb	Ergs per E.M. unit	
Cu → Ag	0	0·137	$0{\cdot}0576 \times 10^6$	2·12
Fe → Constantan	0	3·4	$1{\cdot}42 \times 10^6$	(47·7)
	100	4·1	$1{\cdot}72 \times 10^6$	(50·7)
	200	5·5	$2{\cdot}31 \times 10^6$	(53·7)
Pb → Constantan	0	1·90	$0{\cdot}80 \times 10^6$	27·1
	100	2·73	$1{\cdot}14 \times 10^6$	33·5
	200	3·6	$1{\cdot}51 \times 10^6$	39·9
	300	4·4	$1{\cdot}84 \times 10^6$	46·3
Cd → Pb	0	0·197	$0{\cdot}082 \times 10^6$	3·03
	100	0·390	$0{\cdot}163 \times 10^6$	4·48
	200	0·646	$0{\cdot}271 \times 10^6$	5·93
Cu → Ni	0	1·9	$0{\cdot}80 \times 10^6$	(27)
	100	2·2	$0{\cdot}92 \times 10^6$	(30)
	about 220	2·5 (max.)	$1{\cdot}05 \times 10^6$	(34)
	about 340	1·9 (min.)	$0{\cdot}80 \times 10^6$	(23)
	450	2·4	$1{\cdot}00 \times 10^6$	(25)

In all these cases the Peltier coefficient when expressed in equivalent volts is of the order 10^{-2}—10^{-3} volt. The ratio of the thermal to the electrical conductivity of metals, which contains λ as a factor, shows that $\frac{1}{e}(\lambda_2 - \lambda_1)$ cannot be much greater numerically than this value of P. On the other hand $V_2 - V_1$ may amount to several volts in extreme cases. It follows that $e(V_2 - V_1)$, $w_2 - w_1$ and $\phi_2 - \phi_1$ are approximately equal to one another. Thus $J_m - J_p$ is small compared with $w_m - w_p$ as stated on p. 455.

---

[*] *Loc. cit.* p. 73.

[†] The values in brackets were obtained from specimens of material different from those used in measuring the Peltier coefficient.

The Thomson Effect.

A number of expressions for σ have already been given (p. 451). For the purpose of comparing the results of different theories it is desirable to express some of them in terms of the quantities J etc. which describe the internal structure of the material instead of ϕ etc. Consider, for example, equation (23). We have $\frac{\partial \phi}{\partial T} = \frac{\partial}{\partial T}(W_0 - J)$ and from (2) and (14)

$$\frac{\partial \log \nu_0}{\partial T} = \frac{\partial \log K}{\partial T} + \frac{W_0}{RT^2} - \frac{1}{RT}\frac{\partial W_0}{\partial T} = \frac{W_0}{RT^2} - \frac{J}{RT^2}.$$

Thus

$$\frac{\partial \phi}{\partial T} = \frac{J}{T} - \frac{\partial J}{\partial T} + RT\frac{\partial \log K}{\partial T} = -T\frac{\partial}{\partial T}\left\{\frac{J}{T} - R\log K\right\},$$

so that

$$\sigma = \frac{1}{e}\left\{\frac{R}{\gamma - 1} + T\frac{\partial}{\partial T}\left(\frac{J + \lambda_1 - \lambda_0}{T} - R\log K\right)\right\} \quad \ldots(35).$$

This expression is more general than equation (61) of the last chapter, and the two become identical only in rather simple cases. We have seen that the kinetic theory methods employed in Chapter XVII can only be regarded as strictly accurate when the linear dimensions of the regions, within which the forces exerted on the electrons are appreciable, are small compared with the distance between collisions. In that case the potential energy w of the electrons will only differ by a negligible amount from the mean value J. Also the A of the last chapter is equal to $N\left(\frac{h}{\pi}\right)^{\frac{3}{2}}$ if N is the number of free electrons in unit volume. Since

$$\log N = \log K - \frac{w}{RT} = \log K - \frac{J}{RT},$$

$$\frac{\partial}{\partial T}\left\{\frac{J}{T} - R\log K\right\} = -R\frac{\partial}{\partial T}(\log N) = -R\frac{\partial}{\partial T}(\log A) - \frac{3}{2}\frac{R}{T}.$$

Thus, from (35),

$$\sigma = \frac{R}{e}\left\{\frac{1}{\gamma - 1} - \frac{3}{2} - T\frac{\partial}{\partial T}\left(\log A_1 - \frac{\lambda_1 - \lambda_0}{T}\right)\right\}$$

$$= -\frac{R}{e}T\frac{\partial}{\partial T}\left\{\log A_1 - \frac{\lambda_1 - \lambda_0}{T}\right\}, \text{ if } \gamma = \frac{5}{3}.$$

If the distance between collisions is large compared with the distances within which the forces are appreciable, then the potential energy of the electrons is constant in all but a small fraction of the available space. In this case we can replace the average values of ν and W by actual values without serious error. If, in addition, the law of force during collisions varies as a power of the distance, we see from the last chapter that $\frac{\partial}{\partial T}\left(\frac{\lambda}{T}\right) = 0$. And, finally, if the law of force is the same for the different materials then, for any pair, $\lambda_m - \lambda_p = 0$. Under these circumstances the expressions which have been found reduce to

$$Pe = RT \log \nu_2/\nu_1 \quad \text{.............................(43)},$$

$$\sigma_1 = \frac{R}{e}\left[\frac{1}{\gamma - 1} - T\frac{\partial}{\partial T}(\log \nu_1)\right] \quad \text{...........(44)},$$

$$E = \frac{R}{e}\int_{T_0}^{T} \log\left(\frac{\nu_2}{\nu_1}\right) dT \quad \text{......................(45)},$$

$$\frac{\partial E}{\partial T} = \frac{R}{e}\log\frac{\nu_2}{\nu_1} \quad \text{..............................(46)}.$$

These equations are exact if the collisions are always like those between hard elastic spheres.

Comparison with Experiment.

There are a number of reasons why formulae (43)—(46) are unsatisfactory. In the first place the ratio of thermal and electrical conductivities indicates that the force during collisions does not vary very sharply with the mutual distance. Moreover the computed strength of the centres indicates that the electrons are never free from very considerable forces. Thus the theoretical conditions are far from being satisfied. In the second place the formulae are contradicted by the thermoelectric data themselves. By hypothesis all the electrons are to be treated alike, so that ν in equations (43)—(46) represents the number of current-carrying electrons in unit volume of the material. Now the electrical conductivity is proportional to ν, and there is no reason to expect that the other factors which enter into it will not be of the same order of magnitude for different substances. We should therefore expect that ν_2/ν_1 would, for different substances, be comparable with the ratio of their specific electrical conductivities. Formula (43) would therefore lead us to expect an exceptionally large Peltier effect at the junction between a very good and a very bad conductor. The Peltier effect in a number of cases of this kind has recently been investigated by Koenigsberger and Weiss*. Although these experiments are difficult to make, they seem to have established that the Peltier effect in such cases is not, in general, of exceptional magnitude. In some cases it was found to be in the opposite direction to that given by equation (43). The more complete theory involves two considerations which are neglected in equation (43). In the first place, the law of force during collisions may vary from one material to another. It is, however, extremely improbable that the term $\lambda_1 - \lambda_2$, which arises from this circumstance, can be comparable with the term involving $\log \nu_1/\nu_2$, when the ratio of the two conductivities is very great. We are thus driven to the conclusion that the difference in the potential energy of the electrons is not, in general, measured, even approximately, by $RT \log \nu_1/\nu_2$, where ν_1 and ν_2 are the volume concentrations of the free electrons. This objection cannot be made in the case of the formulae given by the more complete theory, which make the Peltier effect depend on the mean potential energy of all the electrons which may, from time to time, become free under the dynamical actions actually occurring, and not on the actual or average number of those free instantaneously. The number of the former may be quite considerable, although, on account of the intensity of the attraction of the rest of the atom, very few of them are able to get free enough to take part in carrying the current at any instant.

A second point which is of interest in this connection is that raised by the phenomena exhibited at the melting-point. For all the metals which have been tried, except antimony and bismuth, the specific conductivity of the solid at the melting-point is greater than that of the liquid. The changes are quite considerable, the ratio of the two specific conductivities varying from 1·34 in the case of sodium to 4·1 in the case of mercury. In the case of antimony the ratio is 0·70 and in the case of bismuth 0·46. All the metals except antimony and bismuth contract when they solidify. In accordance with (43) and (46) we should expect these changes in the electrical conductivity to be accompanied by very considerable changes in the Peltier effect and the thermoelectric power. Such effects have been looked for most carefully and with negative results. It is clear that the change in the thermoelectric power, if there is any, is much smaller than we should expect from (46). However, it cannot be said that the absence of Peltier effect at the melting-point is properly understood on the more general theory. It is evident that it might happen in a number of ways which it would take too long to discuss here.

---

* *Ann. der Physik*, vol. xxxv. p. 1 (1911).

In the case of the iodiferous cuprous iodide, to which reference has already been made in connection with the Hall effect, Baedeker* finds equation (46) to be accurately verified if the relative values of ν are measured by the relative conductivities. It seems as though the conditions affecting the motion of the electrons are much simpler in the case of this substance than in the case of the metals.

The Conducting Electrons.

The expressions which have been obtained for the various thermoelectric quantities involve integrals which are extended over all the electrons which are able at any time to take part in thermoelectric phenomena. In general we should expect only part of these electrons to be capable of engaging in the transportation of the electric current. Many of them will be so strongly attracted to the atom nearest to them that they will only rarely be able to escape from its immediate neighbourhood. Such electrons will only be slightly displaced under the influence of an external field and will not participate in the conveyance of the conduction current.

It seems natural to suppose that the conduction current is carried by those electrons of the class contemplated which are instantaneously executing open, as opposed to closed, orbits. An expression for the number N of electrons which are executing open orbits may be obtained on the assumption, which accords with the results of the discussion in the last chapter, that the electrons are attracted by centres of force varying inversely as the soluble. In some cases the constituents are only able to dissolve in one another to a definite limited extent. In other cases chemical compounds are formed which give rise to considerable complication. The behaviour in such cases, however, is determined by the mutual solubility of the compounds and the independent constituents; so that the effect of chemical combination is practically to increase the number of substances which have to be considered. Otherwise their behaviour is similar to that of cases where no chemical combination appears to occur.

---

* *Nernst Festschrift*, p. 62, Halle (1912).

Let us consider the two simple cases of complete immiscibility and of complete mutual solubility, from which all the others may be developed. In the case where the constituents are entirely immiscible the specific electrical conductivity, expressed as a function of the composition, changes linearly from the value characteristic of one of the pure substances to that characteristic of the other. Precisely the same statement is true of the thermoelectric power referred to a standard metal. The behaviour of alloys whose constituents are mutually soluble is entirely different in character. The addition of a small quantity of either constituent to the other pure metal produces a large diminution in the specific electrical conductivity. The diminution produced by the addition of a given quantity of the foreign substance diminishes progressively as further amounts are added. Thus the curve which expresses the conductivity as a function of the percentage composition drops sharply from the value corresponding to either pure metal and has a flat minimum in between. The curve which expresses the thermoelectric power as a function of the composition is entirely similar in character. The reader who is interested in the electrical properties of alloys will find a very good account of the recent developments in *Die elektrischen Erscheinungen in metallischen Leitern* by K. Baedeker (Braunschweig, 1911).

We have seen that a considerable change in electrical conductivity is unaccompanied by corresponding changes in the thermoelectric quantities in the case of pure metals at the melting-point. It seems probable that in this case there is something in the conditions of equilibrium which makes $J+\lambda$ take the same value for the electrons in the two phases. This would not necessarily involve a corresponding equality in the fraction in (48). In addition there is the possibility that liquefaction causes magnetite the conductivity increases rapidly with increasing temperature at low temperatures, to a maximum at about 240° C. After that it falls off as the temperature rises in a manner which resembles the behaviour of the metals. The rapid increase at low temperatures is evidently due to the large increase of N with T when ϕ/RT is large. The subsequent decrease may tentatively be attributed to an increase in the number of the repelling centres or to a change in the magnetic structure of the substance. A comprehensive account of the temperature variation of the electrical properties of the comparatively poor conductors, by Koenigsberger, will be found in the *Jahrbuch der Radioaktivität und Electronik*, vol. IV. p. 158 (1907).

The Reflexion of Electrons at the Surface of Conductors.

According to the theory on p. 444 *et seq.*, the concentration ν_0 of the electrons at a point close to a conductor in an enclosure at a constant temperature T is determined by T and the intrinsic potential V_m of the conductor. Thus the equilibrium pressure p has a definite value and in accordance with equation (14) is given by

$$p = \nu_0 RT = ATe^{\int \frac{\phi}{RT^2} dT} \quad \text{.................(51),}$$

where ϕ is the internal latent heat of evaporation reckoned per electron and A is a constant characteristic of the material but independent of T. This equation can be established in a very general manner and is true even if the electrons are emitted wholly or in part by the photoelectric action of the complete aetherial (black body) radiation*. Now in accordance with (28) the number of electrons which reach unit area of the surface of the conductor from outside in unit time is

$$N = \nu_0 \left(\frac{R}{2\pi m}\right)^{\frac{1}{2}} T^{\frac{1}{2}} = A \left(\frac{T}{2\pi mR}\right)^{\frac{1}{2}} e^{\int \frac{\phi}{RT^2} dT} \quad \text{......(52).}$$

In the steady state this quantity must be equal to the number emitted by unit area of the substance in unit time provided all the incident electrons are absorbed by the conductor. However, the experiments of von Baeyer† and the writer‡ have shown that a very considerable proportion even of the very slow moving electrons which are emitted by photoelectric and thermionic action is reflected by metals, so that it is necessary for us to take reflexion into account.

---

* See O. W. Richardson, *Phil. Mag.* vol. XXIII. p. 619 (1912).

† *Ber. der Deutsch. Physik. Ges.* Jahrg. 10, p. 96 (1908).

‡ *Phil. Mag.* vol. XVI. p. 898 (1908), vol. XVIII. p. 694 (1909); *Phys. Rev.* vol. XXIX. p. 557 (1909).

If the body is not a perfect absorber then some of the N electrons will be reflected by it. Let the proportion reflected be r. Then the number actually absorbed by the conductor in time dt is $(1-r)Ndt$. Let the number emitted by the same surface in time dt be ϵNdt. Then the equilibrium condition gives

$$\epsilon = 1 - r \quad \text{..........................(53).}$$

Thus the emissivity (compared with a perfect absorber) and the reflecting power are complementary.

This result is somewhat analogous to Stewart and Kirchhoff's radiation law. By introducing kinetic theory considerations it can be shown that the equality holds for each group of velocities u, v, w, du, dv, dw*. A number of measurements of the reflexion of slow electrons by different metals have recently been made by A. Gehrts†.

Photoelectric Action.

As is well known, the fact, discovered by Hertz, that a spark passes more easily between two terminals when that which is negatively charged is illuminated by ultra-violet light, led to the discovery that the incidence of light of sufficiently high frequency caused the emission of negative electrons from conductors. This phenomenon, which is called the photoelectric effect, is certainly very general and appears to be a universal property of matter. There is no doubt about the universality of this effect when the term light is understood to include X rays and γ rays, although some experimenters have recently called in question the permanence of the effects exhibited by metals and ordinary ultra-violet light‡. We shall now consider what conclusions may be drawn as to the nature of photoelectric action, by the application of the principles of thermodynamics and the kinetic theory of matter.

---

* O. W. Richardson, *Phil. Mag.* vol. XXIII. p. 606 (1912).

† *Ann. der Physik,* vol. XXXVI. p. 995 (1911).

‡ G. Wiedmann and W. Hallwachs, *Ver. d. D. Physik. Ges.* vol. XVI. p. 107 (1914). H. Küstner, *Phys. Zeits.* vol. XV. p. 68 (1914).

Consider an enclosure maintained at the constant temperature T containing a material which is photoelectrically active but which has negligible thermionic emission. No such material may exist; but this will not vitiate the results if thermionic and photoelectric actions are independent, a hypothesis which we shall adopt. The equilibrium concentration of the electrons near the photoelectric material will be determined by equation (51), as we may show by considering the work done against a piston which is transparent to radiation but impervious to electrons. The number which return from the enclosure per unit area of the surface of the material in unit time is therefore given by (52). If α is the proportion of these which is absorbed, i.e. not reflected, it is necessary, in order that the conditions should be steady, that the number emitted in unit time should be equal to

$$\alpha A\left(\frac{T}{2\pi m R}\right)^{\frac{1}{2}} e^{\int \frac{\phi}{RT^2}dT} \quad \ldots\ldots\ldots\ldots\ldots\ldots(54).$$

But the number emitted is a function of the intensity or density of the surrounding radiation. Experiments have shown that for monochromatic radiation the number is almost, if not exactly, proportional to the intensity of the illumination and varies little if at all with the temperature, when the other conditions are constant. We may therefore assume the photoelectric emission to be proportional either to the density of the aetherial radiation present or to its rate of emission or absorption. We shall take the latter as being the more general. Our results can then be adapted to the former hypothesis by making the emissivity ϵ equal to unity. If the steady energy density in the vibrations between ν and $\nu + d\nu$ is $L(\nu)\,d\nu$ the energy belonging to these frequencies which is incident on area dS in unit time is

$$\int \frac{c}{4\pi} L(\nu)\, d\nu\, d\omega\, dS \cos\theta,$$

where $d\omega = 2\pi \sin\theta d\theta$ and the limits are from 0 to $\pi/2$. Thus, if ϵ is the emissivity of the material, the amount of energy of these frequencies which is emitted from unit area in unit time is $\frac{1}{4}c\epsilon L(\nu)\,d\nu$. Let us suppose that the emission or absorption of unit quantity of radiant energy of frequency ν at temperature T

Now consider the average kinetic energy E_ν of the electrons which are emitted by light of frequency ν. The total energy E emitted under the influence of the complete radiation at T is clearly

$$E = \frac{c}{4}\int_0^\infty E_\nu \epsilon F(\nu)\, L(\nu, T)\, d\nu \ldots\ldots\ldots\ldots\ldots(60).$$

This must be balanced by the kinetic energy returned to the metal by the stream of electrons, N per unit time, from outside. If n is the number of electrons per unit volume of the space outside, the stream of energy which reaches unit area in unit time is

$$E' = \frac{n}{2}\left(\frac{km}{\pi}\right)^{\frac{3}{2}} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty um\,(u^2+v^2+w^2)\, e^{-km\,(u^2+v^2+w^2)}\, du\, dv\, dw,$$

and since $k = (2RT)^{-1}$ and $N = n\sqrt{\frac{RT}{2m\pi}}$,

$$E' = 2NRT \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(61).$$

Of this *energy* let the proportion $1-\beta$ be reflected; then β is the proportion which is absorbed. Thus for equilibrium

$$E = \beta \,.\, 2NRT \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(62).$$

If we neglect the effects of reflexion, by putting $\alpha = \beta = 1$, and substitute the value of N given by (47), Chap. XV, making the same approximations as before, we obtain from (60) and (62),

$$\int_0^\infty d\nu\, E_\nu \epsilon F(\nu)\, h\nu^3 e^{-h\nu/RT} = 2A_1 RT^3 e^{-\phi_0/RT} \;\ldots\ldots(63).$$

Subject to

$$\epsilon F(\nu) = 0 \quad \text{from} \quad \nu = 0 \quad \text{to} \quad \nu = \phi_0/h,$$

and $$\epsilon F(\nu) = \frac{A_1 h}{R^2 \nu^2}(1 - \phi_0/h\nu) \quad \text{from} \quad \nu = \phi_0/h \quad \text{to} \quad \nu = \infty,$$

the solution of this is

$$E_\nu = h\nu - \phi_0 \quad \text{when} \quad \phi_0 < h\nu < \infty \;\ldots\ldots\ldots\ldots(64).$$

For values of $h\nu$ which lie between 0 and ϕ_0, E_ν has no meaning, as the corresponding electrons have no external existence.

The solution (64) is dependent on the assumption that the reflexion of electrons can be neglected. This is equivalent to assuming that effects which arise from the collisions of the electrons may be disregarded. But if we consider this from the point of view of the electrons emitted under the influence of the light, we see that neglecting the effect of collisions is tantamount to assuming that the only energy lost by the electrons is used up in overcoming the work ϕ of the forces which tend to retain them in the interior of the substance. Under these circumstances the kinetic energy of the escaped electrons will be equal to that which they acquire by the action of the light (not necessarily from the light directly) minus the work ϕ which they have to do to escape. It is clear that the energy which they acquire under the influence of the light is $h\nu$, where h is Planck's constant. It evidently has the same value for all the electrons liberated by light of the same frequency; any difference in the energy of the electrons emitted by monochromatic light must therefore be attributed to the effect of collisions of the escaping electrons in the interior of the substance.

We can take account of the reflexion of electrons tentatively by putting $\beta = s\alpha$. Then instead of (64) we get

$$E_\nu = s(h\nu - \phi_0) \quad \text{when} \quad \phi_0 < h\nu < \infty \quad \text{.........(65).}$$

For small velocities of incident electrons, such as those with which we are concerned, the proportion reflected increases with increasing energy. It follows that out of a mixed aggregate of incident electrons a greater proportion of the slow ones will be absorbed than of the fast ones, and that the proportion of incident electrons which is absorbed will be greater than the proportion of their incident energy which is absorbed. Thus s will be a positive quantity which is less than unity.

There is no conclusive reason for denying the applicability of the type of argument given above to the emission of material particles of all kinds, whether charged or uncharged and of whatever chemical nature, under the influence of every type of aetherial radiation, provided the chemical actions are of a reversible character so that an equilibrium in the material part of the system can occur.

The foregoing treatment of this subject is taken from papers by the writer*. An equation resembling (64) was first given by Einstein† as a consequence of the view that the energy of light waves was distributed in discrete quanta. For further developments of the subject the reader may consult the following papers:—Einstein, *Ann. der Physik*, vol. XXXVII. p. 832 (1912); *Journal de Physique*, 1913; Planck, *Sitzungsber. d. k. Preuss. Akad. d. Wiss., Math. Phys. Kl.* vol. XVIII. p. 350 (1913); O. W. Richardson, *Phil. Mag.* vol. XXVII. p. 476 (1914).

---

* *Phys. Rev.* vol. xxxiv. p. 146 (1912); *Phil. Mag.* vol. xxiv. p. 570 (1912).

† *Ann. d. Physik*, vol. xvii. p. 145 (1905).

Measurements of the kinetic energy of the electrons emitted by various metals under the influence of light of different frequencies, which have recently been made by the writer and Dr K. T. Compton*, afford considerable support to the above theory. Denoting the maximum observed value of E_ν by E_m, and the mean value by $\overline{E_\nu}$, these quantities were found to satisfy the relations

$$E_m = k_m (\nu - \nu_0)$$

and

$$\overline{E_\nu} = k_\nu (\nu - \nu_0)$$

corresponding to (64) and (65). The values of the constants found are given in the accompanying table:

Metal	Values from E_m			Values from $\overline{E_\nu}$		
	ν_0	λ_0	k_m	ν_0	λ_0	k_ν
Na	51·5	58·3	5·2	52	57·7	2·6
Al	63	47·7	4·3	73	41·1	2·6
Mg	78·5	38·2	5·2	80	37·5	2·55
Zn	80	37·6	5·1	84	35·7	2·8
Sn	83	36·2	4·9	89	33·7	2·75
Bi	91	33	3·55	89	33·7	1·9
Cu	100	30	3·8	97	30·9	1·65
Pt	104	28·8	5·85	103	29·1	2·8

The units are: for ν_0, 10^{13} sec.$^{-1}$: for λ_0, 10^{-6} cm.: and for k_m and k_ν, 10^{-27} erg sec.

According to these results s in (65) is very close to $\frac{1}{2}$ for all the metals investigated. The values of k_m are all somewhat less than the radiation value $h = 6{\cdot}55 \times 10^{-27}$ erg sec., but there are a number of minor causes which might give rise to this discrepancy. There is another way in which h may be estimated from these observations. ν_0, the least frequency which will give rise to any photoelectric emission, is equal to ϕ_0/h. If we may assume that ϕ_0 has the same value as in the thermionic emission and that it is correctly or close to the range covered by the visible spectrum the occurrence of interference, diffraction, refraction and dispersion is taken to indicate that they are aetherial; since it is only on the theory of waves propagated in a continuous medium that the phenomena enumerated have received an adequate explanation. On the other hand if the frequency of the waves were very great it might be extremely difficult, if not impossible, to detect these effects; so that this criterion would not necessarily be available.

---

* *Phil. Mag.* vol. XXIV. p. 575 (1912).

Rather recently W. H. Bragg[*] has suggested a different method of distinguishing between aetherial and uncharged material rays. The distinction depends upon the geometrical distribution of the *secondary* rays which arise when the *primary* rays fall on matter. Recent investigations have shown that these secondary rays may differ greatly in character according to circumstances. Generally speaking, their nature is determined partly by that of the primary radiation and partly by that of the matter on which it impinges. It often happens that the impact of a simple primary radiation on a chemically simple substance will cause the simultaneous emission of more than one type of secondary radiation. The different types of radiation which may thus arise will be considered more fully below. For the present it is only necessary to realize that the secondary rays in general consist partly of negatively charged particles (electronic or β type) and partly of rays which are similar in their properties to Roentgen rays (X type).

Now consider the emission of β secondary rays which occurs when a beam of X primary rays impinges normally on a slab of absorbing matter of indefinite thinness. The number of β rays emitted from the side of the slab on which the primary rays are incident would be expected to be smaller than the number emitted on the emergent side, if the X rays are uncharged material particles. For, in this case, the emitted electrons will be either those which were originally present in the slab and which are knocked out of it by the moving uncharged particles; or they will consist of those of the moving uncharged material particles which have lost a positively charged constituent by collision with the atoms of the slab. In either event the average value of the component of velocity of the emitted electrons along the normal to the surface will not be zero and will be in the direction of the incident radiation.

---

[*] *Nature*, vol. LXXVII. p. 270, Jan. 23 (1908); *Phil. Mag.* vol. XVI. p. 918 (1908).

experienced by this simple theory. In this connection it is worth while to point out that the simple theory does lead to a slight excess of emergent emission owing to the deflexion of the moving electrons caused by the magnetic force in the aetherial pulse. This deflexion, however, is far too small to produce the observed effects if the force in the pulse is supposed to be distributed uniformly, in the usual way. On the other hand the application of thermodynamic and statistical principles to the study of thermal radiation and photoelectric action led us to the view that when radiant energy causes the disruption of an electron from a material system, the electron acquires an amount $h\nu$ of energy, where $h = 6{\cdot}55 \times 10^{-27}$ erg sec. and ν is the frequency of the radiation. This value is in good accord with experiment, although for any moderate value of ν the amount of energy $h\nu$ is much greater than that which we should expect the electron to acquire, on the simple view discussed above, from the direct action of the pulses. It seems fairly clear either that the (electromagnetic) constitution which we have assumed for the radiation is at fault or the mechanism of the process of absorption is different from what we have supposed.

In view of the latter possibility it is very desirable to see if we cannot find out anything about the magnitude of the Bragg and Stuhlmann effects to be expected from aetherial radiations, without making any definite assumption about the way in which the radiation is absorbed, but keeping to the value $h\nu$ of the energy acquired by the disrupted electrons, which, as we have seen, is confirmed by experiments on photoelectric action and by the theory of heat radiation. Consider again the case of aetherial radiation incident normally on a thin slab of absorbing material. In general, absorption may occur through the operation of processes of very different nature, for example, conduction as opposed to accumulation by relatively fixed and stable arrangements; but we shall suppose that the only type of absorption which we need to consider in our slab is that which results finally in the disruption of electrons from material systems. We shall fix our attention on the state of things which exists after the slab has been illuminated for a sufficiently long time so that there is no further accumulation, in the slab, of energy abstracted from the incident beam. Under these circumstances the energy absorbed from the incident radiation will appear, at any rate in the first instance, as the kinetic energy $h\nu$ which the disrupted electrons possess at the instant of disruption. Thus if there are N of them disrupted in unit time the energy absorbed from the radiation is $Nh\nu$. It is important to observe that the energy of the electrons which we have to consider here is that which they possess before, not after, they are emitted from the slab.

But energy is not the only physical quantity of which the incident beam suffers depletion. We have seen in Chap. x, p. 211, that when a material system similar to that under consideration absorbs an amount E of aetherial energy an amount of momentum E/c disappears from the aether. In the present case this momentum must be communicated to the slab and the electrons it contains, in order to satisfy the law of action and reaction. Now according to the electron theory the action of the radiation is on the electrons; so that this momentum is communicated to them in the first instance, and such of it as is ultimately received by the slab of matter only reaches it indirectly through dynamical actions of the nature of collisions. We shall therefore assume that the electrons receive momentum as well as energy from the incident radiation previous to the occurrence of disruption; although we do not know the precise nature of any process which will communicate an appreciable amount either of momentum or energy from a periodic aetherial disturbance to them. Now consider the accumulation, in any small interval of time, of momentum by the electrons in the slab. The increase of momentum of the slab and contained electrons is due to (1) the momentum of electrons which come into the system, (2) the momentum, reckoned negatively, of the disrupted electrons, and (3) the momentum accumulated during the interval by the electrons present in it. Since the state of the electrons instantaneously present in the slab is steady, by hypothesis, it follows that the difference of (1) and (2) is equal to (3). When (1) is zero, (3) is the momentum derived from the radiation. It follows that the momentum which is acquired by all the absorbing electrons from the radiation, is exactly equal to the momentum of the disrupted electrons at the moment of disruption. But since the energy absorbed is $Nh\nu$ the value of the former amount of momentum is $Nh\nu/c$. If $\bar{u}$ is the average component of velocity of the disrupted electrons, in the direction of incidence of the radiation, an have entirely different names depending on more or less fortuitous circumstances. Thus a high speed negative electron is called a β ray if it originates in a radioactive substance, a cathode or Lenard ray if it is produced in a vacuum tube, a secondary Roentgen ray if it is produced by the impact of the Roentgen rays on a solid obstacle, and so on. As it is not our purpose to describe the properties of all the different kinds of radiation in detail we shall extend the scope of the term β ray so as to cover any negatively charged material ray. In the same way we should call a positively charged material ray an α ray, although this term also is usually applied only to those rays which originate with radioactive substances. The uncharged, and presumably aetherial, radiations will be referred to as X rays. We shall use the term γ rays for these when they originate from radioactive substances and Roentgen rays when they are produced in vacuum tubes. For a detailed account of the properties of these radiations the reader may be referred to the following authorities: J. J. Thomson, *Conduction of Electricity through Gases*, Chaps. XI, XII, XIX and XX; Rutherford, *Radioactivity, passim.*

The difference between α and β rays is not merely one of sign. In all cases the specific charge (e/m) has been found to be of a different order of magnitude. For the β rays e/m always has the large value which corresponds to electrons, whereas for α rays e/m is always of the same order as the corresponding quantity in electrolysis. As all the evidence points to the charge e being either equal to, or a small multiple of, the elementary electronic charge in all these cases, it follows that the α rays consist of atoms or molecules which have lost one or more negative electrons. Their properties do not furnish any evidence that we have succeeded in isolating any fundamental positively charged electrical atom which would correspond to the negative electron, unless it be the hydrogen atom which has lost a negative electron.

The most convenient test for the presence of the various types of radiation under consideration is that furnished by the production of electrical decomposition (ionization). Thus the passage of the rays through an insulating gas imparts to it the property of electrical conductivity. The amount of this ionization is often taken to be a quantitative measure of the energy absorbed from the rays. In the case of the X radiations this has proved to be the forces called into play by the resulting displacement of the electron from its original position of equilibrium. The limitations thus introduced will be considered below. The energy radiated by the electron during the complete passage of the pulse is evidently

$$\int dW = \frac{e^4}{6\pi m^2 c^3} \int \frac{X^2 dt}{(1-\beta^2)^3} \text{ ....(4)},$$

where the integrals are extended over the time of passage. It is probable that, except in the case of very penetrating radiations, we shall not be led into serious error if we neglect β^2 compared with unity. Thus

$$\int \frac{X^2 dt}{(1-\beta^2)^3} = E\frac{d}{c} = \frac{W}{c} \text{ ....(5)},$$

where $E = \overline{\tfrac{1}{2}(X^2 + H^2)} = \overline{X^2}$ is the energy in unit volume of the primary pulse and d is its thickness. If N is the number of electrons in unit volume of the matter, and if the primary pulse only loses energy in this way, the energy $-\delta W$ which is lost by the primary pulse in travelling a distance δz is given by

$$\delta W = \delta(Ed) = -\frac{Ne^4}{6\pi m^2 c^4} W \delta z \text{ ....(6)}.$$

Thus the relation between the energy W of unit area of the pulse and the distance z traversed in the matter is

$$W = W_0 e^{-\frac{Ne^4}{6\pi m^2 c^4} z} \text{ ....(7)}$$

if W_0 is the value of W when $z = 0$.

On the theory that we are considering, the thickness of a primary pulse is clearly equal to the distance traversed by an electromagnetic disturbance in the time during which the acceleration of the emitting electron is appreciable. This time will be determined by two factors: (1) the velocity of the moving electron relative to the retarding atom, and (2) the geometrical distribution of the field of force, inside the atom, which produces the acceleration. The velocity of the moving electrons is always large and has the value c of the radiation velocity as an upper limit. The linear dimensions of the fields of force will depend upon the constitution of the atom, but will be comparable with the distance between the electrons which are present in the latter. As we the distance*. When the velocity V of the β particles is so large that the resulting deflexion is small, the deflexion produced by a single encounter is $\frac{2e^2}{mV^2}\frac{1}{x}$, where x is the perpendicular distance between the electron encountered and the undeflected path. The average value of this for all the electrons which lie within a distance a of the line of motion is $\frac{4e^2}{mV^2}\frac{1}{a}$. If l is the average length of the path of the moving particle which lies inside the atom, if the electrons are uniformly distributed in the atom and if there are ν of them per unit volume of the latter, the number of collisions with electrons which lie within a distance a is $\nu\pi a^2 l$. If we suppose for the moment that the negative electrons are alone operative and consider any very large number σ of encounters with atoms the average total deviation would be

$$\frac{4e^2}{mV^2}\frac{1}{a}\sqrt{\sigma\nu\pi a^2 l}=\frac{4e^2}{mV^2}\sqrt{\sigma\nu\pi l}.$$

Since this must be $\sqrt{\sigma}$ times the effect of a single encounter the effect of the latter is

$$\theta_1=\frac{4e^2}{mV^2}\sqrt{\nu\pi l}=\frac{16}{5}\frac{e}{mV^2}\frac{1}{b}\sqrt{\frac{3N_0}{2}} \quad\ldots\ldots\ldots\ldots(9),$$

substituting $\sqrt{l}=\frac{4}{5}\sqrt{2b}$, where b is the radius of an atom.

If ϕ_1 is the average deflexion due to the sphere of positive electrification the theory of central forces shows that this is given by

$$\phi_1=\frac{e^2}{mV^2}\frac{N_0}{b}\frac{\pi}{4} \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10),$$

provided ϕ_1 is small. When the positive electricity is made up of definite units the mean deflexion ϕ_2 due to these is given by

$$\phi_2=\frac{16}{5}\frac{e^2}{mV^2}\frac{1}{b}\sqrt{\frac{3N_0}{2}}\sqrt{1-\left(1-\frac{\pi}{8}\right)\tau^{\frac{1}{3}}} \quad\ldots\ldots(11),$$

where τ is the ratio of the volume occupied by the positive electricity to the volume of the atom.

Thus, according as we adopt the first or second hypothesis as to the geometrical configuration of the positive electricity, the value of θ will be given by the equations

$$\theta=(\theta_1^2+\phi_1^2)^{\frac{1}{2}} \text{ or } \theta=(\theta_1^2+\phi_2^2)^{\frac{1}{2}} \ldots\ldots\ldots\ldots(12),$$

---

* Routh, *Dynamics of a Particle*, Cambridge, 1898, Chap. VI.

and the average deflexion ψ_m in passing through a thin plate of thickness t will be given by either

$$\psi_m = \frac{e^2}{mV^2}\left\{\frac{384}{25}N_0 + \frac{\pi^2}{16}N_0^2\right\}^{\frac{1}{2}}\sqrt{N\pi t}\ \dots\dots\dots\dots\dots(13),$$

or

$$\psi_m = \frac{e^2}{mV^2}\left\{\frac{384}{25}N_0\left[2 - \left(1 - \frac{\pi}{8}\right)\tau^{\frac{1}{3}}\right]\right\}^{\frac{1}{2}}\sqrt{N\pi t}\ \dots(14).$$

As the only unknown quantities on the right-hand sides of (13) and (14) are N_0 and τ, if the equations are supported by experiment they ought to enable us to determine the value of N_0 the number of electrons in the atom.

The extent to which these equations are in accordance with the observed scattering of β rays when they pass through thin sheets of various solids has been examined by Crowther*. If ψ is any particular angle of deflexion it follows from the theory of errors that the probability of a deflexion greater than ψ is $e^{-\psi^2/\psi_m^2}$. Thus from (8) the thickness t_0 for which this probability is $\frac{1}{2}$ will be given by

$$e^{-\psi^2/ct_0\theta^2} = \tfrac{1}{2}\ \dots\dots\dots\dots\dots\dots\dots\dots(15),$$

or

$$\psi/t_0^{\frac{1}{2}} = \theta\sqrt{(c\log 2)}\dots\dots\dots\dots\dots\dots\dots(16),$$

where c is a constant for any particular substance. The probability that the deflexion is less than ψ is equal to $1 - e^{-\psi^2/ct\theta^2}$, that is to

$$1 - e^{-k/t}\ \dots\dots\dots\dots\dots\dots\dots\dots\dots(17),$$

where k is constant for any particular substance, if ψ is kept constant. From (8) and (16)

$$\psi/t_0^{\frac{1}{2}} = \psi_m/t^{\frac{1}{2}} \times \sqrt{\log 2}\ \dots\dots\dots\dots\dots\dots(18),$$

and from either (13) or (14)

$$\psi/t_0^{\frac{1}{2}} = \text{const.} \times \frac{1}{mV^2}\ \dots\dots\dots\dots\dots\dots(19),$$

where the absorbing medium is kept the same and the velocity of the incident rays is varied. When ψ is also kept constant

$$mV^2/t_0^{\frac{1}{2}} = \text{constant}\dots\dots\dots\dots\dots\dots\dots(20).$$

The method of investigation adopted by Crowther was to measure the proportion of the incident β rays whose deflexions were less than a fixed value ψ when they passed through absorbing sheets of different materials and when the thickness of see below that there is another criterion which may be used instead of k, namely the maximum velocity of the β rays emitted when the X rays fall on metals. This has been shown to have, for each pure radiation, a definite value which is independent of the metal (cf. also p. 510).

---

* *Roy. Soc. Proc.* A, vol. LXXXIV. p. 226 (1910).

When the primary rays are such as are emitted by an ordinary Roentgen ray bulb the difference between the secondary rays emitted by aluminium (atomic weight 27) and an element like copper (atomic weight 63) is very striking. In the first place, the quantity of the secondary emission (scattered radiation) from aluminium is comparatively insignificant. The amount of the secondary radiation from the copper is very much larger. The two radiations differ also in quality. We have seen that the aluminium radiation has the same penetrating power as the primary radiation. The copper radiation is much more absorbable. It is also very nearly a pure radiation. The current evidence is to the effect that it consists almost entirely of a pure radiation mixed with a trace of scattered radiation. The amount of the scattered radiation from copper, and other elements of atomic weight above 40, appears to be of the same order as that given out by aluminium.

This pure secondary radiation is that which we have referred to as the characteristic secondary radiation. It is characteristic in the sense that its absorbability k has a value which is characteristic of the metal from which it is emitted. In the case of the elements whose atomic weights lie between about 40 and 100 the absorbability of the characteristic secondary rays is quite independent of the nature of the primary X rays which are used to excite them. With the elements of still higher atomic weight the phenomena are more complicated and it appears that these elements give off more than one kind of characteristic radiation. It is convenient to allude to the characteristic radiation from a particular element M as the M-X rays. Thus the characteristic rays from copper are called the Cu-X rays. This notation is due to Bragg.

The absorbability of the characteristic rays diminishes, and their penetrating power increases, as the atomic weight of the parent element increases. The characteristic rays differ from the scattered rays in one important particular in addition to those already

(Fig. 52). The abscissae are proportional to the absorption coefficient of the various radiations in aluminium. The experimental values of k/ρ in aluminium are indicated by the abscissae corresponding to the various vertical dotted lines alongside which are written the chemical symbols of the different elements by which the pure characteristic radiations were emitted. The ordinates are the values of k/ρ, where ρ is the density of the absorbing element, for the elements C, Mg, Fe, Ni, Cu, Zn, Ag, Sn, Pt and Au. In the case of the elements of atomic weight below 40, namely C and Mg, the relation between the absorbability of the radiations by the respective elements and that by aluminium is a linear one, as has already been pointed out on p. 497. The same thing, however, will be observed to be true in the case of Pt*, Au*, Ag and Sn as well. In fact such a linear relation between the absorption coefficients for different pure radiations has been found to hold quite generally, provided the range of radiations tested neither includes that characteristic of the absorbing substance nor lies near the less absorbable side of it.

The curves for Fe, Ni, Cu and Zn are quite different. As they all exhibit the same features we need only consider one of them, that for nickel for example. The absorbability by nickel of the radiations from Cr, Fe, Co and Ni, which are either more absorbable than, or as absorbable as, that characteristic of nickel, is proportional to the absorption of the same radiations in aluminium. The next most penetrating radiation, as measured on the aluminium scale, is the Cu-X radiation, and this shows a small but definite increase over the absorbability required by the law of proportionality. As the absorbability in terms of aluminium of the characteristic rays diminishes further, their absorbability in nickel increases very rapidly. Thus the absorption coefficient of the Zn-X rays in nickel is several times that of the Cu-X rays, although the atomic weight of zinc (65) is only 2 units greater than that of copper. Further diminution in the aluminium absorbability of the characteristic rays is accompanied by an approximately proportional diminution in the absorbability of the rays in nickel. Thus the curve again approximates to a straight line which has a different slope from the linear portion on the more absorbable side of the radiation which is characteristic of nickel. It will be noticed that in each case the radiation having the maximum absorbability is that which is characteristic of the element next but one to the absorbing element when they are arranged in the order of increasing atomic weight.

---

* It is necessary to except the penetrating Ag-X and Se-X rays in these cases.

Direct experiments have shown that the abnormal absorption of radiations which are slightly more penetrating than the radiation characteristic of the substance tested, is the cause of the emission of the characteristic radiation. The *amount* of the characteristic radiation emitted by an element is found to be zero until the exciting rays become at least as penetrating as the characteristic radiation; as the penetration, measured in terms of aluminium or some substance which behaves similarly to aluminium, is increased beyond this critical value, the amount of characteristic radiation which the exciting radiation will cause to be emitted, compared with the ionization which it produces in air or any other gas containing only light atoms, increases at first to a maximum value and then diminishes. In fact, the quantity of characteristic radiation emitted appears, for all exciting radiations, to be proportional to the excess of absorption over what would be given by the linear relation which holds for exciting radiations that are more absorbable than the characteristic radiation.

Since these results appear to be quite general, the occurrence of humps, like those shown in the relative absorbability curves for Zn, Cu, Ni and Fe in Fig. 52, affords a simple and delicate method of detecting the presence of unknown characteristic radiations. Working in this way Barkla and Nicol* have succeeded in separating two quite distinct characteristic radiations from each of the elements silver, antimony, iodine and barium. A comprehensive examination of all the radiations thus far discovered shows that, when they are arranged in the order of k/ρ, their comparative absorbability in aluminium, they fall into two series. These have been called by Barkla the K and L series respectively. In each series the order of diminishing absorbability is that of increasing atomic weight of the metal of origin. The data for the radiations which have thus far been isolated are given in the following table:—

---

* *Nature*, Aug. 4 (1910).

Radiation	Atomic Weight	Absorbability of Fluorescent Radiation (k/ρ in Al)		Remarks
		Series K	Series L	
H-Mg	1·008—24·32			No radiation observed. Probably very absorbable.
Al and S	27·1 and 32·07			Radiation observed but k/ρ not yet measured.
Ca	40·09	435		Less accurate than succeeding values.
Cr	52	136		The value in brackets is deduced by Barkla from X ray measurements. The value for Rb has not been determined so accurately as the others.
Fe	55·85	88·5		
Co	58·97	71·6		
Ni	58·68 (61·3)	59·1		
Cu	63·57	47·7		
Zn	65·37	39·4		
As	74·96	22·5		
Se	79·2	18·9		
Br	79·92	16·4		
Rb	85·45	13·7		
Sr	87·62	9·4		
Mo	96	4·7		
Rh	102·9	3·1		
Ag	107·88	2·5	700	Values for L more accurate than for K series.
Sn	119	1·57		
Sb	120·2	1·21	435	
I	126·92	0·92	306	
Ba	137·37	0·8	224	
Ce	140·25	0·6		
W	184		33	These values are comparatively approximate.
Pt	195		27·5	
Au	197·2		25	
Pb	207·1		20	
Bi	208		19	

Reasons have been given by Barkla for believing in the existence of series other than the series K, L which have already been explored*. It will be noticed that the fact that the characteristic radiations are only excited by radiations more penetrating than themselves is analogous to Stokes's law in optics: according to which fluorescent light is invariably of lower frequency than the light which excites it. For this reason the characteristic radiations have sometimes been called fluorescent radiations. It is obvious that each pure characteristic radiation has features which are that they are able to travel in a gas like air before they cease to produce additional ionization of the gas. This distance has been found to be sufficiently definite to be taken as a satisfactory index of the velocity of the quickest rays in any given group.

---

* *Phil. Mag.* vol. XXII. p. 408 (1911).

Using this criterion Sadler* and Beatty† have been able to show that the maximum velocity of the β rays emitted when a given characteristic X radiation is allowed to fall on different metals always has the same value. This maximum velocity is therefore characteristic of the X radiation and independent of the metal; it is greater the greater the atomic weight of the metal of which the X rays are characteristic.

This question has been studied more fully by Whiddington‡. He has shown by experiment that if the original velocity of a group of β rays is v_0, their velocity v after traversing a thickness d of matter is given by

$$v_0^4 - v^4 = ad \qquad \text{.......................(25)},$$

where a is a constant characteristic of the matter. Since (25) appears to hold with fair approximation down to the velocity (approximately $v = 0$) at which ionization ceases, it follows that the velocity v_s of the fastest secondary rays is given by

$$v_s^4 = ad \qquad \text{.........................(26)},$$

where $a = 2 \times 10^{40}$, for air at atmospheric pressure, and d is the maximum distance in which ionization is perceptible. Analysing the results of Sadler and Beatty in a manner which depends on this principle it appears that

$$v_s = k'w \qquad \text{.........................(27)},$$

where k' is a constant and w is the atomic weight of the radiator of which the secondary rays are characteristic. This relation has been verified for the Fe-, Cu-, Zn-, As-, Sn-, Mo-, and Ag-X rays. The value of k' is very close to $10^8 (\pm 4\,\%)$.

We have seen that the characteristic X rays are only excited by X rays more penetrating than themselves. Using an X ray tube with a silver anticathode Whiddington§ has investigated the relation between the velocity of the cathode rays in the primary electrons emitted is proportional to the incident intensity, whereas the maximum energy of the electrons is independent of this intensity. These facts receive a simple and obvious explanation on the view that the X rays and light consist of showers of material particles or of bundles of energy which are localized in space and do not spread out as the distance from the source is increased. On the other hand in the case of light (and also X rays, see p. 507) it is difficult to account for the phenomena of interference and refraction on such a view. It seems a little safer, therefore, to suppose tentatively that the energy of the emitted electrons in some way represents a condition which determines the disruption of matter under the stimulus of a given radiation. To agree with the results of the theory of black body radiation and of the experiments on the emission of electrons under the influence of light it is necessary that one part of this condition should be that the energy of the disrupted electrons is either equal to $h\nu$, where $h = 6{\cdot}55 \times 10^{-27}$ erg sec. and ν is the frequency of the radiation, or is an integral multiple of this quantity. It also appears that this condition must be of a very general character and necessarily inherent in all types of matter.

---

* *Phil. Mag.* vol. XIX. p. 337 (1910).

† *Phil. Mag.* vol. XX. p. 320 (1910).

‡ *Roy. Soc. Proc.* A, vol. LXXXVI. pp. 360, 370 (1912).

§ *Roy. Soc. Proc.* A, vol. LXXXV. p. 323 (1911).

We have seen that the maximum energy of the electrons emitted under the influence of illumination by light is given by the equation

$$T_m = \nu h - w_0 \quad \text{......................}(31),$$

where w_0 is a constant characteristic of the type of matter by which the light is absorbed. If this relation holds for the radiations of the X type as well as light, the results given above enable us to determine the frequency of the characteristic X radiations. In these cases w_0 is negligible compared with νh, so that

$$\nu = \frac{mv_s^2}{2h} = \frac{mk'^2(Aw+B)^2}{2h} \quad \text{..............}(32).$$

Thus for the copper-X rays belonging to the K series, putting

$$k' = 10^8,\ A = 1,\ w = 63,\ B = 0,\ m = \frac{1{\cdot}6}{1{\cdot}77} \times 10^{-27} \text{ and } h = 6{\cdot}55 \times 10^{-27},$$

$$\nu = 2{\cdot}74 \times 10^{18}\ \text{sec.}^{-1}.$$

According to this estimate their frequency is about 2000 times as great as that of visible light.

series and the diffuse subsidiary series is equal, for each component, to the frequency of the first corresponding term of the diffuse subsidiary series.

Roughly speaking the intensity of the lines tends to diminish as m increases, although the relative intensities vary a great deal with the mode of excitation of the lines. In a sharp principal series of doublets the component of higher frequency is more intense than the slower component. The order of variation of intensity of the components with increasing frequency is usually the reverse of this in the case of the lines belonging to either of the subsidiary series.

The most important steps in the establishment of the serial relations just described are due to the work of Runge[*], Rydberg[†], Kayser, Paschen and Bergmann. These relations have been somewhat generalized recently by Ritz and Paschen and the results used, following a suggestion of Rydberg, to predict the existence of "combined" series. The success which has attended this development may be judged from papers by Paschen in the *Annalen der Physik* from 1908 to the present time[‡].

The X-ray spectra of the elements have been described in the preceding chapter, p. 510.

Series in Band Spectra.

A number of relations between the emission frequencies observable in band spectra have been discovered by Deslandres[§]. The frequencies of the lines which make up a single band are found to agree very approximately with the formula

$$\nu = A(m+a)^2 + c \quad \text{.....................(18)},$$

where A, a and c are constants and m takes various successive integral values. As a rule this formula does not deviate seriously from the observed values unless the band contains more than about 50 lines.

In general, emission bands are not found isolated but occur in related groups and in some cases there are a number of groups is very small in comparison with either, we can write, instead of (26) and (27),

---

[*] *Brit. Assoc. Report for* 1888, p. 576.

[†] *K. Svenska. Vet. Akad. Handl.* vol. XXIII. No. 11, p. 155 (1890).

[‡] See also Stark, *Principien der Atomdynamik*, vol. II. p. 50.

[§] Papers in *Comptes Rendus* from 1885–1891 and from 1902–1905.

$$\delta n_1 = n_1 - n_0 = \frac{eH_z}{2mc}, \quad \delta n_2 = n_2 - n_0 = -\frac{eH_z}{2mc} \quad \ldots(28).$$

Thus the difference between the frequency of each of the two circular vibrations and that of the original vibration has the same magnitude but is positive in the one case and negative in the other. The frequency difference per unit magnetic field is quite independent of the frequency of the original line and is entirely determined by the universal electronic constant e/m.

Now consider the radiation which will be emitted by an electron moving in the manner we are contemplating. Since the circular vibrations are capable of resolution into simple harmonic motions, the nature of the radiation, although not necessarily its amount, can be deduced from the simpler case of a rectilinear simple harmonic motion. By Poynting's theorem the rate of transmission of energy at any point is proportional to the vector product of the electric and magnetic intensities and the direction of transmission is the direction of that vector. It follows from the results of Chapter XII that in any motion in which the velocity and acceleration are collinear the magnetic intensity lies in circles about the axis of motion and vanishes at points along this axis. There is therefore no radiation along the direction of the axis of a simple harmonic motion. Now turn to the radiation in any of the directions which are perpendicular to this axis. It follows from Chapter XII that the electric intensity in the radiation wave in this case lies in the plane containing the axis of motion and the radius. It is also perpendicular to the latter. The magnetic intensity is equal to the electric intensity and its direction is normal both to the electric intensity and the radius. Since the plane of polarization of the radiation is that which contains the direction of the magnetic intensity in the wave-front, we see that the radiation emitted in the direction under consideration is completely polarized in the equatorial plane. These principles together with equations (23)—(25) are sufficient to determine completely the character of the radiation which is emitted in directions coincident with, and normal to, the lines of magnetic force.

Consider the light emitted along the lines of force first. The simple harmonic motion given by (25) emits nothing in this

To illustrate this it will be best to consider one or two examples. The behaviour of the lines in the sharp principal series of doublets is exemplified by the case of the sodium D lines. Their behaviour in a transverse magnetic field is exhibited in Fig. 53. The letters p and n denote that the lines are polarized parallel and perpendicular to the lines of magnetic force, respectively.

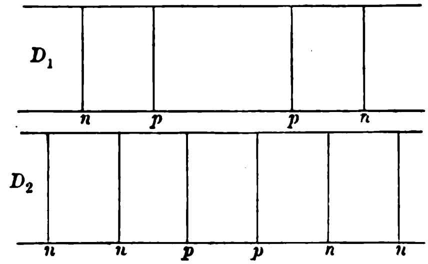


Fig. 53.

The two components of sharp subsidiary series of doublets are decomposed in exactly the same way except that the less refrangible component of the one series replaces the more refrangible of the other and vice versa. This may be regarded as a confirmation of the correspondence in the structure of the lines of these two series which is indicated by the respective series formulae.

The strong lines of the diffuse subsidiary series of doublets are split up into triplets, whilst the satellites split into eight components of which six are polarized perpendicularly and two parallel to the axis of the magnetic field.

The way in which the lines of the sharp subsidiary series of triplets break up in a transverse magnetic field is shown in Fig. 54. The behaviour of the diffuse series of triplets is still more complex, as is also that of many lines which have not been assigned to any series.

Paschen and Back* have recently observed that with certain double lines the character of the Zeeman effect depends very much on the strength of the magnetic field. The arrangement of the lines in strong fields is in some respects simpler than that in weak fields.

Not much has yet been accomplished in the way of an explanation of the more complicated types of Zeeman effect, although observed in the direction of the magnetic field, gave rise to components with circular polarization in the opposite direction to that exhibited by the line spectra. Other lines showed the normal behaviour in this respect and in both cases the polarization was incomplete. The existence of circular polarization in the opposite direction to that given by the normal effect cannot necessarily be interpreted as implying positively charged vibrators, since Lorentz* has shown that this effect can arise from the vibrations of negative electrons under suitable circumstances.

---

* *Ann. der Physik*, vol. xxxix. p. 897 (1912), vol. xl. p. 960 (1913).

Theories of Spectral Emission.

The frequency of the occurrence of the normal type of Zeeman effect and the fact that the more complicated types are closely related to the simple type and to the series of line spectra, show that the oscillations of negative electrons play a very important part in spectral emission. The chief difficulty which now lies in the way of the further development of the interpretation of these effects arises from our ignorance of the nature of the emitting systems. If we consider a particular series of lines, for example, we can either regard these as the overtones which accompany the fundamental vibration of a single system or we can look upon each line as the natural vibration of a separate system. In the latter case the other lines of the series are attributed to the occurrence of systems which are constitutionally related to one another in some regular way. Both of these views have their respective advocates. Another view, which attributed all the spectral lines to the natural modes of vibration of the normal atom, is certainly no longer tenable. As each electron can at most only give rise to three natural frequencies the number of electrons per atom which would be required to furnish the spectra of elements like iron and titanium is quite prohibitive. Moreover such a hypothesis would make the absorption spectra of metallic vapours quite different from what they are.

The experimenters who have adopted what one may call the overtone view of the nature of spectral series, have come to quite different conclusions as to the nature of the emitting systems. Thus Lenard† found that the upward stream of colour arising from a bead of alkali salt in the outer regions of a flame is undeflected by an electric field, whereas in the interior of the flame the coloration is deflected. In the former case only the lines of the principal series are emitted, whereas in the latter case lines of the subsidiary series may also be found. He therefore concluded that the emitters of the principal series were uncharged atoms of the metal and those of the subsidiary series atoms which had lost one or more electrons. Experiments with salt vapours in the arc confirmed this conclusion. On the other hand Stark*, largely from experiments made on the rapidly moving positive ions (canal rays) found in vacuum tubes under certain circumstances, has come to the conclusion that the emitters of all the line series are positive ions: and holds, on other grounds, that the band spectra are emitted by the neutral particles. There is no doubt that the sources of the line spectra are in many cases in rapid motion, since Stark† has shown that they exhibit the Doppler effect. This fact alone does not settle the question, but there is no doubt that the streams of deflected positively charged particles do in general give rise to the emission of series lines. However, such streams usually contain a fair proportion of neutralized particles which might be the source of the emission. In fact it seems to the writer that the bulk of the experimental evidence which has been brought to bear on this question might be interpreted in various ways.

---

* *Theory of Electrons*, p. 123.

† *Ann. der Physik*, vol. XI. p. 636 (1903), vol. XVII p. 195 (1905).

The truth of the overtone view of the nature of spectral lines cannot be said to have been established. In fact there are very grave objections to it. R. Ladenburg‡, who investigated the dispersion of luminous hydrogen in a vacuum tube, found that it was negligible except in the neighbourhood of the red line H_α. In this region satisfactory measurements showing the regular type of anomalous dispersion to be expected near a natural period were obtained, although no measurable effect could be detected in the neighbourhood of the bright blue-green line H_β. An extensive series of measurements covering the dispersion of the vapours of the different alkali metals in the neighbourhood of the lines of the various principal series has recently been carried out by Bevan§.

---

* *Jahr. der Radioakt. u. Elektronik*, vol. VIII. p. 231 (1911).

† *Phys. Zeits.* vol. VI. p. 892 (1905).

‡ *Ver. der Deutsch. Physik. Ges.* x. Jahrg. p. 858 (1908).

§ *Roy. Soc. Proc.* A, vol. LXXXIV. p. 209 (1910), vol. LXXXV. p. 54 (1911).

two lines of Balmer's series, so that this argument leads to the conclusion that the systems which can emit the different lines of this series are not present in equal numbers. It follows that the different lines of a given series are given out by different systems and therefore presumably by atoms in different states.

Bevan's results with the lines of the sharp principal series of the alkali metals confirm these conclusions. The argument here is not so strong because these lines do not exhibit the normal type of Zeeman effect, so that we have not such good grounds for the validity of the simple theory of dispersion which leads to (29). The constants might be of the more general type given in Chap. VIII, p. 176, for example. However, even in this case, it is not likely that the order of magnitude of ν_s estimated from the simple formula would prove to be seriously wrong.

As to the actual numerical differences in ν_s Bevan finds for the first four doublets of the principal series of potassium that the numbers of vibrating electrons are proportional to the respective numbers:

$$m_1 = 0{\cdot}113,\ m_2 = 1{\cdot}58 \times 10^{-4},\ m_3 = 1{\cdot}1 \times 10^{-5} \text{ and } m_4 = 3 \times 10^{-6}.$$

There is some evidence that the value of ν_s is equal, or approximately equal, for the two lines of a doublet. Results of a similar character are given by the other alkali metals. The proportion of systems capable of emission to the total number of atoms present is probably greatest in the case of the centres which emit the sodium D lines. Using data given by Wood for the temperature 644° C. Bevan estimates that the proportion of centres to atoms is about 1 in 12. A corresponding calculation has been carried out by Ladenburg and Loria in the case of hydrogen in a Geissler tube. By making use of the formulae (24 a) and (24 b) of Chap. VIII, in which the part played by the absorption term is taken into account, they conclude that about one centre capable of emitting the line H_α is found in every 50,000 molecules, under the conditions of their experiment.

As to the nature of the difference between the different atoms which makes them capable of emitting different spectral lines a number of plausible hypotheses may be considered. In the first place the electromagnetic analysis (measurement of the deflexion produced by transverse electric and magnetic fields) has shown that the atoms present in the vapours of metals may lose varying numbers of negative electrons. Thus Sir J. J. Thomson* has recently found that the positive rays in mercury vapour contain mercury atoms which have lost 1, 2, 3, 4, 5, 6 and 7 electrons respectively. In the second place there have also been shown to occur in vacuum tube discharges atoms which have gained one or more electrons in excess of the normal amount. Finally, systems may be formed by the combination of two or more atoms. I ought also to add that Ritz† has shown that if the electrons are supposed to vibrate in the field of a small magnet, different frequencies occur which obey the same kind of law as Rydberg's series, if the magnet is supposed to be built up of varying numbers of smaller elementary magnets. The possibility of this last type of theory has received some support from Weiss's work on the magnetic properties of bodies considered in Chap. XVI, but it does not otherwise seem to agree very well with present tendencies in the development of the theory of atomic structure (see Chap. XXI). H. A. Wilson‡ has pointed out that if we take the atom to be composed of a number of electrons in equilibrium inside a sphere of positive electrification of uniform volume density—a hypothesis which, as we shall see in the next chapter, gives a fair account of many of the properties of the atoms—then each atom possesses only one mode of vibration which is effective in producing any considerable amount of radiation, and the frequency of this mode of vibration depends only on the universal constants e, m and c, and the density ρ of the positive electrification in the sphere. According to this theory the frequencies of all spectral lines are determined by the density ρ of the positive electrification. For a given atom the requisite changes in ρ are secured either by a deficiency in the normal complement of electrons or by combination of the atoms with each other. As these changes always take place in discrete amounts the frequencies can be made to depend on whole numbers in the same way as in Balmer's series.

A number of attempts have been made to construct vibrating systems which have overtones resembling the spectral series.

---

* *Phil. Mag.* vol. XXIV. p. 668 (1912).

† *Gesammelte Werke*, Paris (1911).

‡ *Phil. Mag.* vol. XXIII. p. 660 (1912).

Ritz[*] has succeeded in constructing two-dimensional dynamical systems of this kind, whilst Whittaker[†] has expressed the form which the requisite energy functions must take in a very general manner.

A theory of spectral emission, due to Bohr, which depends on different principles from any of the foregoing will be considered in the next chapter, p. 585. This theory offers a quantitative explanation of Balmer's series and is able to overcome most of the difficulties we have so far encountered.

Fluorescence.

Many substances when illuminated, let us say by monochromatic light, are found to emit light of a different colour or frequency. This phenomenon is called fluorescence. Sir George Stokes, who made very important researches on the subject, concluded that the fluorescent light was invariably of lower frequency than the exciting light. This generalization has been found to be only approximately true. The first exceptions to it were noticed by Lommel[‡]. The more recent investigations of Nichols and Merritt[§] have shown that in the case of certain substances which they examined, the relative distribution of energy in the fluorescent spectrum does not depend very much on the frequency of the exciting light, and that a considerable proportion of the emitted energy may belong to higher frequencies than that of the exciting radiation.

Very often the fluorescent emission lasts for some time after the exciting radiation has been cut off. The phenomenon is then often termed phosphorescence. There does not seem to be any very sharp line of demarcation between fluorescence and phosphorescence, as the duration referred to may take almost any value from zero upwards with different substances and according to circumstances. However, in the case of liquid and gaseous bodies this duration has always been found to be too small to this hypothesis seems wide enough to cover all the phenomena referred to, and so far as the writer is able to judge, there are no facts known at present which definitely contradict it.

---

[*] *Loc. cit.*

[†] *Roy. Soc. Proc.* A, vol. LXXXV. p. 262 (1911).

[‡] *Wied. Ann.* vol. III. p. 113 (1878).

[§] *Phys. Rev.* vol. XIX. p. 18 (1904).

Wiedemann and Schmidt* discovered that the vapours of various organic compounds and of the metals sodium and potassium exhibited fluorescence. This list has been extended to include the additional elementary substances mercury, iodine and bromine by R. W. Wood. The fluorescence exhibited by the vapours of the alkali metals and iodine has been investigated in detail by Wood† and found to exhibit remarkable results. Although there is a good deal of similarity in the effects exhibited by the different substances, perhaps the case of iodine is most interesting. When the stimulation is effected by monochromatic light the fluorescent spectrum is not a continuous one but is in the form of a series of fine lines whose frequencies are approximately equidistant, with the original line forming one of the lines of the series. The positions of the lines of the fluorescent spectrum change in a remarkable manner as the frequency of the exciting light is altered. They are also intimately related to the fine structure of the very complex absorption spectrum of iodine. This spectrum has seven very fine lines, visible only with the highest resolving power, within the width of the green mercury line. When the green mercury line is used for excitation it is found that a large number of the fluorescent lines are made up of fine lines having a structure similar to the seven absorption lines covered by the exciting spectrum, only the fluorescent lines are about thirty times as far apart on the frequency scale as the corresponding absorption lines. Many other interesting peculiarities have been observed which promise important results when investigated further.

The line fluorescence shown by the elements alluded to does not obey Stokes's law. It possesses the remarkable property‡ of being converted into a band spectrum by an admixture of inert gases like helium. The intensity of the banded fluorescence diminishes with the amount and molecular weight of the strange gas added. No adequate theory of these phenomena has yet been put forward, although it is clear that any satisfactory theory of for left-handed circularly polarized light is shifted towards the low frequency side. In the direct Zeeman effect we found that when e was negative the component of lower frequency was left-handed, if we observed the emission in the direction of the magnetic force.

---

* *Ann. der Phys.* vol. LVII. p. 447 (1896).

† Wood's *Physical Optics*, chap. XVIII., New York (1905); *Phil. Mag.* Oct. 1912.

‡ Wood and Franck, *Ver. der Deutsch. Phys. Ges.* vol. XIII. p. 84 (1911).

Effects of the kind just specified were first observed by Maculoso and Corbino[*] in their experiments on the absorption of light by sodium vapour in the neighbourhood of the D lines, in the presence of a magnetic field. More complete experiments by Zeeman[†] and Hallo[‡] have shown that the phenomena are completely in accordance with the theory, the development of which is largely due to Drude and Voigt. J. Becquerel[§] has shown that similar phenomena are displayed by the very sharp absorption bands exhibited by the salts of certain rare earth metals at low temperatures, whilst Wood[||] has found them in the fine lines of the channelled absorption spectrum of sodium vapour. In both these cases the direction of rotation for some of the lines corresponds to that given by the elementary theory for positively charged particles (see p. 524).

We shall now turn to the case in which the light is propagated at right angles to the lines of force of the external magnetic field which we shall still suppose to lie along the z axis. The equations of motion of the electrons will therefore still be given by equations (33)—(35). Let the light be propagated along the x axis. The various electric and magnetic vectors will then depend upon the coordinates only through the factor $e^{\iota p(t-qx)}$. In the first place we notice that when the light is plane polarized so that the electric vibration is in the same direction as the magnetic field, the latter exerts no effect on the observed phenomena. For the relations between the polarization, the electric intensity and the motion of the electrons are given by equation (35) together with

$$D_z = P_z + E_z \quad \text{and} \quad P_z = Nez.$$

These equations, together with

$$\frac{\partial H_y}{\partial x} = \frac{1}{c}\frac{\partial D_z}{\partial t} \quad \text{and} \quad \frac{\partial E_z}{\partial x} = \frac{1}{c}\frac{\partial H_y}{\partial t},$$

---

[*] *Comptes Rendus*, vol. cxxvii. p. 548 (1898).

[†] *Amsterdam Proc.* vol. v. p. 41 (1902); *Arch. Néerl.* (2) vol. vii. p. 465 (1902).

[‡] *Arch. Néerl.* (2) vol. x. p. 148 (1905).

[§] *Comptes Rendus* (1906–7).

[||] *Phil. Mag.* vol. xv. p. 274 (1908).

corresponding lines in the longitudinal inverse effect or the corresponding emission lines in the direct effect.

To the same order of accuracy the maximum coefficient of absorption is evidently, from (57), the same for both lines and is given by

$$\frac{kp}{c} = \frac{p}{4\delta c} \quad \text{.........................(61).}$$

This is just one-half of the value of the maximum absorption coefficient in the absence of a magnetic field under otherwise identical conditions (see Chap. VIII, p. 167). Thus when the medium is traversed perpendicularly to the direction of the magnetic force, by light in which the electric intensity is also perpendicular to the magnetic force, the light will be absorbed to the extent given by (61) when the frequency has either of the values given by (60). Corresponding to one of these vibrations the polarization of the medium and the motion of the electrons will be given respectively by $P_x = +iP_y$ and $x = +iy$, and corresponding to the other by $P_x = -iP_y$ and $x = -iy$. Thus the electrons move in opposite directions in circles perpendicular to the direction of the magnetic force. This motion agrees with that which corresponds to light of the same frequency when the inverse longitudinal Zeeman effect is under consideration. In general these transverse effects will not be symmetrical about the frequency corresponding to $\gamma = 0$ because the expressions for n and k involve odd as well as even powers of γ.

Magnetic Rotation and Double Refraction.

The rotation of the plane of polarization of plane polarized light, when it passes through a refracting medium in the direction of the lines of force of an applied magnetic field, was discovered by Faraday* when he was engaged in trying to find the connection between electromagnetism and light. The explanation of this, the first magneto-optical effect to be discovered, follows at once from the theory of the longitudinal inverse Zeeman effect which has just been given.

As is well known, a plane polarized ray of light is equivalent to two rays circularly polarized in opposite senses and differing variables. By introducing the condition that the charge of the electrons is to be the same in both cases we obtain an additional equation between k and m, but there is still nothing to determine the absolute value of these quantities. Thus, corresponding to any solution of the equations, there are an infinite number of other solutions corresponding to linear transformations of the variables. If any solution corresponds to a state of steady motion or rest there are an infinite number of such possible states. Thus the finite size and general definiteness of the atom must be due to something outside the equations referred to.

---

* *Exp. Res.* § 2152 (1845).

This conclusion is valid only if the dimensions of the elements of electric charge are negligible compared with the dimensions of the problem. The radius of an electron, as deduced from its inertia, is about 10^{-13} cm. and is therefore about 10^{-5} times that of an atom. We should therefore expect that, as an approximation towards the problem of atomic structure, we may treat the negative electron as a point charge endowed with inertia. In that case there must be something which determines the finiteness and definiteness of the atom, which lies outside the properties of the negative electron itself as well as outside the equations of the field. It may therefore be expected to depend upon the properties of the positive electricity. One way in which this additional requirement has been met is by supposing the positive electrification to be limited to the region inside a sphere of atomic dimensions. This hypothesis was first introduced by Lord Kelvin[*]. Its consequences have been worked out very fully by Sir J. J. Thomson[†] and the question has also been considered by Lord Rayleigh[‡]. Other possibilities will be referred to later. (See pp. 580—587.)

A number of writers, in addition to those referred to above and in the preceding chapter, have considered the question of the emission of series of spectral lines by atoms made up of electrons. Among these are Jeans[§], Nagaoka[||] and Schott[¶]. One of the by vibrations about these positions, their definiteness is accounted for. On the other hand the existence of paramagnetism leads us to infer the presence, at least in certain atoms, of electrons rapidly revolving in closed orbits. Consequently numerous attempts have been made to seek the origin of spectral lines in the kinematics of revolving rings of electrons. Unless rather special and somewhat strange assumptions are made, however, it is impossible to secure sufficient definiteness of frequency in this way.

---

[*] *Phil. Mag.* vol. III. p. 257 (1902).

[†] *Phil. Mag.* vol. VII. p. 237 (1904).

[‡] *Phil. Mag.* vol. XI. p. 117 (1906).

[§] *Phil. Mag.* vol. II. p. 421 (1901); vol. XI. p. 604 (1906).

[||] *Phil. Mag.* vol. VII. p. 445 (1904).

[¶] *Phil. Mag.* vol. XII. p. 21 (1906); vol. XIII. p. 189 (1907); vol. XV. p. 438 (1908).

The difficulty here, which has also been emphasized by Jeans, is brought out in the following discussion which is due to Schott*. Consider a system containing a ring whose radius is r consisting of n electrons, evenly spaced about the circumference when undisturbed, and revolving with velocity $V = \beta c$. Approximate stability is assured by the presence of a central positive charge or a containing sphere of positive volume electrification or some other device giving rise to an attraction towards the centre of the ring. The equations of motion are shown to be

$$nU = -\frac{r^2}{c^2}\frac{\partial (mV)}{\partial t} \quad \ldots\ldots\ldots\ldots\ldots\ldots(1),$$

and
$$P_1 - \frac{c^2}{r^2}(1+\beta^2)K + \frac{e^2}{r^2}W^2 = \frac{mV^2}{r} \ldots\ldots\ldots\ldots(2),$$

where U is a function of n and β which is proportional to the rate of radiation from the ring†. The terms in K and W in (2) represent the force which arises from the electrons in the ring itself, other than the one whose motion is considered. K is a function of n only and W of n and β. P_1 is the central force arising from the rest of the atom and the term on the right is the centrifugal force. P_1 may be considered as a function of r into whose form it is not necessary to inquire. m the mass of an electron is of the form $\frac{e^2}{c^2 a}\psi(\beta)$, where a is its radius. Now nU, being proportional to the radiation, is always positive if the system radiates at all. Thus if a is constant equation (1) shows that β cannot be constant for a radiating system. When β is given r may be obtained from (2), so that the radius of the ring also will depend on their orientation in space*. When bodies whose atoms contain systems of electrons in orbital motion are turned, a gyrostatic couple ought to be experienced. The fact that such an effect has not been observed cannot, I think, be regarded as disproving the existence of such systems. The moment of momentum of the revolving electrons is so small that one could only expect to detect such an effect by carrying out special experiments of a delicate character, and this has not yet been done.

---

* *Phil. Mag.* vol. XII. p. 21 (1906).

† Cf. the value of U in *Phil. Mag.* vol. XII. p. 22 (1906) and of R in vol. XIII. p. 194 (1907).

In the papers of Schott to which we have referred, the questions of the stability of, and the amount of radiation from, rings of revolving electrons are both considered in detail. In most of the calculations the only assumption made as to the nature of the force attracting the electrons to the centre of the atom is that it is of the usual electromagnetic type. Some of the results are therefore more general than the corresponding ones given by the investigation which we shall now describe.

The properties of the type of atom which is made up of coplanar rings of electrons inside a large sphere of uniform indeformable positive electrification have been worked out by Sir J. J. Thomson† and lead to very interesting results. If b is the radius of the large sphere and ρ the density of the electrification in it, then the force, acting on an electron whose charge is e at a distance a from the centre of the sphere and due to the positive charge of the latter, is $\frac{4}{3}\pi\rho e a$. If there are ν electrons in the neutral atom, the charge E in the positive sphere is

$$E = \nu e = 4\pi\rho b^3/3.$$

Hence, in terms of the number and charge of the electrons, the force they experience at a distance a from the centre due to the positive sphere is $\nu e^2 a/b^3$. In addition to this the electrons are acted upon by forces due to their mutual repulsions. If O is the centre of the sphere an electron at A will be acted upon by an electron at B with a force e^2/AB^2 and the radial component of this $= \frac{e^2}{AB^2}\cos OAB$. If $OA = OB$, i.e. if the electrons are in the same ring, this repulsive force is $e^2/4OA^2 \sin \frac{1}{2}AOB$. Hence, if we have n electrons arranged at equal angular distances $2\pi/n$ round

---

* Cf. Jeans, *Phil. Mag.* vol. XI. p. 606 (1906).

† *Loc. cit.*, cf. also *The Corpuscular Theory of Matter*, chap. VI., New York, 1907.

The radial repulsion exerted by the sth electron on the pth is

$$R_{ps} = -e^2 \frac{\partial}{\partial r_p}\left[r_p^2 + r_s^2 - 2r_p r_s \cos(\theta_s - \theta_p) + (z_p - z_s)^2\right]^{-\frac{1}{2}}$$

$$= \frac{e^2}{4a^2 \sin\psi}\left\{1 - \frac{\rho_p}{a}\left(\tfrac{3}{2} - \frac{1}{2\sin^2\psi}\right) - \frac{\rho_s}{a}\left(\tfrac{1}{2} + \frac{1}{2\sin^2\psi}\right) - \tfrac{1}{2}(\phi_s - \phi_p)\cot\psi\right\} \quad \ldots(6),$$

where $\psi = (p - s)\dfrac{\pi}{n}$.

The tangential force Θ_{ps} tending to increase θ_p is

$$\Theta_{ps} = -\frac{e^2}{r_p}\frac{\partial}{\partial\theta_p}\left(\frac{1}{\left[r_p^2 + r_s^2 - 2r_p r_s \cos(\theta_s - \theta_p) + (z_p - z_s)^2\right]^{\frac{1}{2}}}\right)$$

$$= \frac{e^2}{4a^2}\frac{\cos\psi}{\sin^2\psi}\left\{1 - \tfrac{3}{2}\frac{\rho_p}{a} - \tfrac{1}{2}\frac{\rho_s}{a} - (\phi_s - \phi_p)(\cot\psi + \tfrac{1}{2}\tan\psi)\right\} \quad (7).$$

The force perpendicular to the plane of the undisturbed orbit is

$$Z_{ps} = \frac{e^2}{8a^3 \sin^3\psi}(z_p - z_s) \quad \ldots\ldots\ldots\ldots\ldots\ldots(8).$$

In equations (6)—(8) higher powers than the first of the small quantities ρ, ϕ and z have been dropped.

The radial force R_p exerted by all the other electrons in the ring on the pth is

$$R_p = \frac{e^2}{4a^2}S - \rho_p A' - \Sigma_s \rho_{p+s} A_s - a\Sigma_s \phi_{p+s} B_s \quad \ldots\ldots(9).$$

The tangential and perpendicular forces on the same particle are respectively

$$\Theta_p = \Sigma_s \rho_{p+s} B_s - a\phi_p C + a\Sigma_s \phi_{p+s} C_s \quad \ldots\ldots\ldots(10),$$

and

$$Z_p = z_p D - \Sigma_s z_{p+s} D_s \quad \ldots\ldots\ldots\ldots\ldots\ldots(11),$$

where

$$\left.\begin{aligned}
&S = \sum_{s=1}^{n-1}\left(1\Big/\sin\frac{s\pi}{n}\right), \quad A' = \frac{e^2}{4a^3}\left(\tfrac{3}{2}S - \tfrac{1}{2}\sum_{s=1}^{n-1}\left[1\Big/\sin^3\frac{s\pi}{n}\right]\right),\\
&A_s = \frac{e^2}{8a^3}\left(\frac{1}{\sin\frac{s\pi}{n}} + \frac{1}{\sin^3\frac{s\pi}{n}}\right), \quad B_s = \frac{e^2}{8a^3}\left(\cos\frac{s\pi}{n}\Big/\sin^2\frac{s\pi}{n}\right),\\
&C_s = \frac{e^2}{4a^3}\frac{\cos\frac{s\pi}{n}}{\sin^3\frac{s\pi}{n}}\left(\cot\frac{s\pi}{n} + \tfrac{1}{2}\tan\frac{s\pi}{n}\right), \quad C = \sum_{s=1}^{n-1} C_s,\\
&D_s = \frac{e^3}{8a^3}\frac{1}{\sin^3\frac{s\pi}{n}}, \text{ and } D = \sum_{s=1}^{n-1} D_s.
\end{aligned}\right\} \quad (12).$$

values of the frequency. When $k=0$, $M_k=0$ and equation (22) reduces to a quadratic, so that altogether there are $4(n+1)/2-2=2n$ values of the frequency. When n is even there are $n/2+1$ equations, but $M_k=0$ when $k=0$ or $n/2$, so that two of these reduce to quadratics. Thus we see that whether n be odd or even there are $2n$ possible frequencies for vibrations in the plane of the orbit, corresponding to the $2n$ degrees of freedom of the electrons, in this plane.

In a similar way the n frequencies q of the vibrations at right angles to the plane of the orbit may be shown to be given by

$$\frac{\nu e^2}{b^3}+P_k-P_0-mq^2=0 \text{ ....}(24),$$

where
$$P_k=\frac{e^2}{8a^3}\sum_{s=1}^{n-1}\cos\frac{2sk\pi}{n}\Big/\sin^3\frac{s\pi}{n} \text{ ....}(25).$$

The values of L, M, N and P have been worked out in a number of the simpler cases and the corresponding frequencies calculated[*].

The importance of this investigation lies, for reasons which have already been given, not so much in the fact that it enables us to evaluate the frequencies of the vibrations of the electrons, as in the fact that the frequency equations (22) and (24) enable us to determine whether the equilibrium is stable or not. If all the values of q given by (22) and (24) are real then all the disturbances ρ, ϕ and z are periodic functions of the time, and the system will only execute small oscillations about the steady configuration. In this case the steady motion is stable. If, on the other hand, any of the roots contain an imaginary part, the corresponding values of ρ, ϕ or z will contain factors of the type e^{ht}, where h is a real positive constant. Thus any such disturbance will increase indefinitely with the lapse of time and the presence of complex or imaginary values of q shows that the equilibrium is unstable. The condition that the equilibrium should be stable is therefore that all the values of q given by the equations (22) and (24) should be real for all the admissible values of k.

---

[*] Cf. Thomson, *loc. cit.*

distributed, is five[*]. Six or more electrons in a single ring are never in stable equilibrium however great the angular velocity is. Such a ring may, however, be made stable if a negative charge is introduced at the centre. If this negative charge is equal to that carried by p electrons, the effect of the additional repulsion thus introduced is to increase A in the frequency equation (27) to $A + 3\frac{pe^2}{a^3}$[*]. This can always be made positive by taking p sufficiently large.

From what has been said it is clear that the condition for a ring to be stable is that $A + 3\frac{pe^2}{a^3} = \frac{3}{4}\frac{e^2S}{a^3} - (L_0 - L_k) + 3\frac{pe^2}{a^3}$ should be positive. The greatest value of $L_0 - L_k$ is for $k = n/2$ when n is even and for $k = (n-1)/2$ when n is odd. The number of electrons which would have to be placed at the centre of the ring to ensure stability is therefore given by the least integral value of p for which

$$\left.\begin{aligned} &3\frac{pe^2}{a^3} > L_0 - L_{\frac{n}{2}} - \tfrac{3}{4}\frac{e^2S}{a^3} \text{ when } n \text{ is even} \\ \text{and}\quad &2\frac{pe^2}{a^3} > L_0 - L_{\frac{n-1}{2}} - \tfrac{3}{4}\frac{e^2S}{a^3} \text{ when } n \text{ is odd}\end{aligned}\right\} \quad \ldots (28).$$

By means of these equations the least value of p for a given value of n can be calculated. p increases very rapidly for large values of n as is shown by the following table of corresponding values:

n	5	6	7	8	9	10	15	20	30	40
p	0	1	1	1	2	3	15	39	101	232

When p exceeds unity the additional electrons cannot all lie at the centre of the ring but must separate under their mutual repulsion. Thus when $n = 9$ and $p = 2$ there will be an external ring of seven and two electrons inside situated along a diameter of the sphere at equal distances from the centre. When p is large it is possible for the electrons to arrange themselves in stable equilibrium in rings rotating about a common axis and lying in parallel planes. On the assumption that the effect of external rings of electrons can be neglected and that the effect of internal rings is the same as if the total charge on the electrons were number in the centre increases to four. With seventeen electrons it is found that the stable arrangement is an outer ring of eleven with six inside it. We have seen, however, that a ring of six is unstable, so that one of them goes to the centre leaving a ring of five. We have thus three concentric systems in this case, containing respectively one, five and eleven electrons. When the number of electrons in the outer ring becomes considerable, the number which it is necessary to place inside the ring in order to ensure stability increases very rapidly. For large values of the number n in the ring the number p inside varies as n^3.

---

[*] Cf. Thomson, *loc. cit.*

The actual arrangement of the electrons with some of the smaller numbers is given in the following table. The number of horizontal columns gives the number of rings; the numbers are the number of electrons in each ring. The total number in the atom is thus obtained by adding up the vertical columns.

outer ring	1	2	3	4	5

outer ring	5	6	7	8	8	8	9	10	10	10	11
inner ring	1	1	1	1	2	3	3	3	4	5	5

outer ring	11	11	11	12	12	12	12	13	13	13	13	14	14	15	15
2nd ring	5	6	7	7	8	8	8	8	9	10	10	10	10	10	11
1st ring	1	1	1	1	1	2	3	3	3	3	4	4	5	5	5

outer ring	15	15	15	16	16	16	16	16	16	16	17	17	17	17	17	17	17
3rd ring	11	11	11	11	12	12	12	13	13	13	13	13	13	14	14	15	15
2nd ring	5	6	7	7	7	8	8	8	8	9	9	10	10	10	10	10	11
1st ring	1	1	1	1	1	1	2	2	3	3	3	3	4	4	5	5	5

outer ring	17......	19	20	20	20	20	20	20	20	20	20	21	21
4th ring	15......	16	16	16	16	17	17	17	17	17	17	17	17
3rd ring	11......	13	13	13	13	13	13	13	14	14	15	15	15
2nd ring	5......	8	8	8	9	9	10	10	10	10	10	10	11
1st ring	1......	2	2	3	3	3	3	4	4	5	5	5	5

outer ring	21............	22............	23............	24
5th ring	17............	20............	20............	21
4th ring	15............	16............	17............	17
3rd ring	11............	13............	13............	15
2nd ring	5............	8............	10............	11
1st ring	1............	2............	3............	5

Thus 93 electrons arrange themselves in six concentric rings of 5, 11, 15, 17, 21 and 24 respectively. 94 electrons begin the series with seven rings having the arrangement 1, 5, 11, 15, 17, 21 and 24.

Numbers like those in the preceding table showing the arrangement of the electron in the atom can be obtained in the following manner. It is first necessary to determine a sufficient number of corresponding values of p and n. This can be done by means of equations (28). Having done this we can draw the curve which gives p as a function of n. It is of the general form exhibited by the accompanying figure.

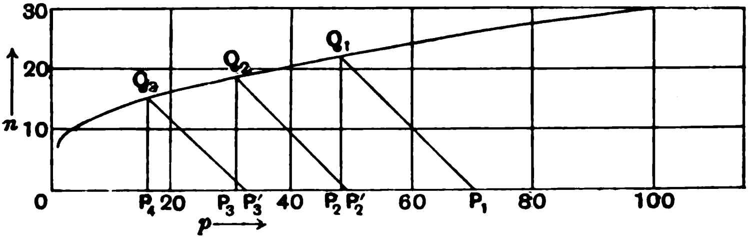


Fig. 56.

Suppose it is required to find how a large number N of electrons will arrange themselves in stable rings. Take $OP_1 = N$ along the axis of p and draw P_1Q_1 inclined at 135° to OP_1 intersecting the curve in Q_1. Then if Q_1P_2 is perpendicular to OP_1, $Q_1P_2 = P_2P_1$. By the property of the curve Q_1P_2 electrons in a ring require OP_2 inside to make them stable, so that $N = OP_1$ electrons will arrange themselves with an outer ring of P_2P_1 surrounding OP_2 electrons. In general OP_2 will not be an integer; in this case, instead of OP_2 we take OP_2' the next higher integral value. Thus we see that $N = OP_1$ electrons will arrange themselves so that the outermost ring contains $P_2'P_1$ electrons. We can find the distribution of the OP_2' electrons by repeating the process. From P_2' draw $P_2'Q_2$ inclined at 135° to OP_1 and draw Q_2P_3 perpendicular to OP_1. If OP_3' is the next integer higher than OP_3 the OP_2' electrons will arrange themselves with an outer ring of $P_3'P_2'$ electrons surrounding the remaining OP_3'. In this way we can proceed until the whole N electrons are accounted for.

Referring to the table on the preceding page we see that the successive atoms formed in this way possess features analogous to those properties of the chemical elements which are summarized by the periodic law. Thus it will be observed that the first element with seven rings is the same as that with six except two electrons before the number falls to 58 and it becomes exceedingly stable as regards further emission. Thus it will resemble the divalent electropositive elements of the alkaline earth group. In a similar way the next element will not so easily emit an electron but it will be able to part with three before reaching the very stable condition with 58 electrons. It will therefore be less electropositive but will be capable of giving rise to a trivalent positive ion. It will therefore resemble the trivalent earth metals such as aluminium. In this way we see that as the number of electrons increases, the elements which have an outer ring of twenty become continuously less electropositive but have a continuously increasing electropositive valency. When we come to the last atom with twenty electrons in the outer ring and 67 altogether this would theoretically be able to emit nine electrons, but practically it will be so stable that it will be incapable of emitting any except under very great forces. It may thus be considered to resemble the inert gases helium, argon, etc. which are incapable of entering into chemical combination. Again, this element will be unable to combine with a free electron; for if it did so it would have 68 electrons altogether and these would arrange themselves with an outer ring of 21. As the system with 68 electrons and an outer ring of 21 is very unstable and liable to emit an electron when neutral, it will be still more unstable when it carries an excess of negative charge, so that this atom with 67 electrons will be incapable of combining with one additional electron. It will thus have zero electronegative valency. In this respect it again resembles the inert gaseous elements. The element with 66 electrons will tend to combine with one electron; since, when the atoms are neutral, the element with 67 is more stable than that with 66. It will not be able to combine with more than one, for if it did it would possess 68 electrons altogether, an arrangement which, as we have seen, is exceedingly unstable even when neutral. This element will thus behave like the strongly electronegative monovalent elements fluorine, chlorine, etc. The atom with 65 electrons will have a less strongly marked tendency to combine with an additional one but will be able to combine with two altogether before reaching the limiting condition. It will thus resemble the elements oxygen, sulphur, etc. which are divalent but less strongly electronegative

It is worth while remarking that the elements which lie between the very strongly electronegative monovalent elements like fluorine and the next higher strongly electropositive monovalent element, e.g. sodium, are the gases He, Ne, Ar, Kr and Xe which have no chemical affinity; whereas those which lie between elements like manganese, which show some analogy with the chlorine group, and the corresponding element like copper, which is analogous to sodium, exhibit the highest electropositive valency of any of the elements. This is shown by the oxy-compounds and various complex amino-compounds of the metals of the iron, palladium and platinum groups.

It is convenient to have a name for the measure of the tendency of an electron to leave an atom which has been discussed above. Following Sir J. J. Thomson we shall refer to this as the electronic pressure of the atom.

Chemical Combination.

It has been pointed out that Lord Kelvin, in his paper entitled "Aepinus atomized" (*Phil. Mag.* vol. III. p. 257, 1902), was the first to suggest that the chemical atom consists of a sphere of uniform positive electrification containing negative electrons of much smaller dimensions embedded in it. In that paper he discusses the forces which will come into play when some of the simpler types of atoms are brought together. Considering the simplest type of all, that which contains only one electron, it is evident that two such atoms will exert no mutual force if the spheres lie entirely outside one another. If, however, one of the spheres A penetrates another B, then, since part of the positive sphere of B lies inside that of A, the repulsion of the positive sphere of B by that of A will be less than the attraction of the negative electron at the centre of A. Thus the two spheres will attract one another. It is clear, however, that the negative electron at the centre of each sphere will still be in equilibrium there until the centre of one sphere lies within the circumference of the other. If both the spheres are equal, the mutual repulsions of the negative electrons then exceed their attractions by the positive spheres; the electrons therefore move along the line joining the centres of the two atoms so as to lie outside the centres, but remain always within the atom. We thus get a to that between an uncharged conductor and a charged sphere. This effect will take place whether the atoms are similar or dissimilar.

The combination of two atoms in this way will, in general, be accompanied by a transference of electrons from one to the other. The way in which this takes place is most conveniently described in terms of the idea of electronic pressure. Consider the atom which possesses N_2+1 electrons. Let us denote this by A and suppose it to be brought into immediate juxtaposition with the atom B, which contains N_3-1 electrons. We have seen that the electronic pressure of the atom A is very high while that of B is very low. In other words less work will be required to drag an electron out of A than out of B. Under these circumstances we should expect an electron to pass from A to B; so that in the compound thus formed A will carry one electronic unit of positive and B one electronic unit of negative electricity. After this transference has taken place A will contain N_2 electrons so that its electronic pressure will be very low. Thus no more electrons will pass over to B, especially since if another were to be transferred B would then have N_3+1 electrons, and this number corresponds to an atom with a very high electronic pressure. There is still another reason why no further transference should take place, and that is that the transference of the first electron produces an electric field between the atoms which tends to stop any further exchange.

If another atom of A were brought into the neighbourhood of the molecule AB it would not be able to transfer another electron to B, despite its own high electronic pressure. For if it did so B would then have N_3+1 electrons, an arrangement which is particularly likely to shoot off one electron. This action would certainly result in this case since it would be helped by the electric field from B to A. Thus A is clearly a monovalent electropositive element and B is a monovalent electronegative element. The molecule AB is a fully saturated molecule and will show no tendency to enter into further combination with any other elements.

Now consider the interaction between an atom C containing N_2+2 electrons and an atom D of another element containing N_3-2 electrons. The electronic pressure of C being higher than that of D, we should expect an electron to pass from C to D.

field. These forces would not be so intense as those which come into play during chemical combination proper, but they are probably capable of accounting for such phenomena as hydration, cohesion, surface tension, latent heat and the properties arising from molecular forces generally.

In considering the nature of the bonds connecting atoms of the same element it is interesting to observe that the phenomenon of self-combination is only shown to any considerable degree by two elements, carbon and silicon, and that both these elements lie exactly in the centre of their respective series in the periodic table. Their electrochemical properties are therefore neither markedly electropositive nor electronegative, and they ought to show an almost equal tendency to enter into combination in either sense. This is just the kind of condition of the atom that we should expect would give rise to self-combination, on the view that we have adopted as to the mechanism of this action. For such an atom would be almost equally stable whether it formed the positive or the negative end of a chemical bond. It would show a greater degree of adaptability to the effect of the remaining groups with which the atoms were combined than would atoms which were not so constituted.

If this view of chemical combination is the correct one, it becomes a matter of great importance to determine the sign and magnitude of the charge carried by each atom in different compounds. The most generally applicable method is, of course, the electrolytic one. In the case where one atom is deposited at the cathode it presumably carries a positive charge in the compound; since, so far as the writer is aware, no cases are known where the solvent is capable of reversing the polarity of the ions formed by a given electrolyte.

The phenomenon of magnetism also seems to be capable of shedding some light on this phenomenon. Townsend has shown that the magnetic permeability of solutions of salts of iron containing the same amount of iron has the same value for all ferric salts; it also has the same value, but one which is different from the preceding, for all ferrous salts, whilst for the ferricyanides it has uniformly the value zero, i.e. the ferricyanides are non-magnetic. These results show that the magnetism of iron compounds is an atomic property but indicate that it depends on combine with the haloid acids are readily interpreted on this view, as also are the chemical properties of the diazo-compounds and a number of other substances.

It is necessary to add that most of these facts have been accounted for by chemists in other ways which seem fairly satisfactory.

Reviewing the whole question broadly it seems quite likely that in a great number of cases of chemical combination transference of electrons between the atoms will not occur. There seems to be little doubt that the forces between uncharged atoms can be sufficiently great to account for the energy of chemical combination in a great number of cases. Definiteness of valency can also be accounted for in this way. For instance, if the neutral carbon atom possesses four electrons arranged at the corners of a regular tetrahedron, the directions of maximum electric intensity in the field of the neutral atom will be along lines possessing a similar tetrahedral symmetry. The safest course to adopt at present would appear to be that of restricting the interpretation of valency bonds as representing electronic transference to those cases only in which the possibility of electrolytic dissociation has been demonstrated.

The Structure of the Positive Electricity.

The foregoing, necessarily brief, review of chemical phenomena shows that there is a very close correspondence between the properties of the elements and those required by the atoms considered by Thomson. It is not likely that the hypothesis of a sphere of positive electrification of a uniform volume density is essential in order to arrive at conclusions of the same general character as those which have been indicated. It is probable that somewhat similar conclusions about questions of stability would hold if the volume density of the positive electrification, instead of being uniform, were greatest at the centre of the sphere. On the other hand, we have seen that it is impossible to construct a definite atom out of indefinitely small elements of positive and negative electrification acting on each other according to the so that the greater part of the mass is entirely unaccounted for by this theory.

The following hypothesis appears to offer a possibility of explaining these facts and at the same time of retaining the important features of the positive sphere. Suppose that the positive electricity, instead of being uniformly distributed, is in the form of electrons whose charge is E and mass M, where E is very small compared with the numerical value of the charge e of the negative electrons. Since the electromagnetic mass of such particles is proportional to E^2/R, where R^3 is comparable with their volume, it is clear that if R were small enough the positive electricity could be made to carry most of the inertia of the atom. If $e = \nu E$, then $r/R = \nu e/m \div E/M$, where the small letters refer to the negative electrons. If the hydrogen atom contains only one negative electron, then E/M is the value of the corresponding quantity for the hydrogen atom in electrolysis. r would be greater than R in the proportion of something comparable with 1000ν, so that the positive electron would have to be confined to a much smaller space even than the negative. This would lead to the difficulty about the definiteness of the atom already alluded to, if the law of force between these positive electrons were that of the inverse square. Let us suppose that at very small distances this law does not hold but is replaced by something more complicated, let us say

$$+\frac{a}{r^2} - \frac{b}{r^p} + \frac{c}{r^{p_2}},$$

where $p_2 > p_1 > 2$. At very small distances the third term would give a repulsion and keep the positive electrons from joining together, and the first term would give the usual law of force at large distances. The middle term would cause the positive electrons to attract one another at certain distances. This would make them aggregate into clusters which, if the constants were of suitable magnitude, would be of dimensions comparable with that of the atoms. The positive electrons would be regularly distributed inside so that such clusters would behave very much like a continuous distribution of positive electrification, provided E were sufficiently small.

On this hypothesis there would be a definite positive atom capable of existing without the presence of a negative electron.

The Radioactive Elements and the Periodic Law.

The study of the chemical properties of the radioactive elements[*] has brought to light a number of facts of the highest importance which bear on the relation between the chemical properties of the elements and their atomic weights. Soddy pointed out that when a radioactive element *A* was converted into a second element *B* with accompanying emission of an α ray, then the chemical properties of *B* would be those of an element in the next column but one before *A* in the periodic table. Russell[†] showed that if a β ray were emitted instead of an α ray the new element would be found in the next column beyond that of the first element. Thus, to illustrate the case by considering one particular chemical property, the emission of an α ray diminishes the electropositive valency by two, whereas the emission of a β ray increases it by one. In the case of the thorium series for example, if we neglect branch products, the successive changes are exhibited in the following table:

$$\underset{\text{IV}}{\text{Th}} \xrightarrow{\alpha} \underset{\text{II}}{\text{Msth I}} \xrightarrow{\text{Rayless}} \underset{\text{III}}{\text{Msth II}} \xrightarrow{\beta} \underset{\text{IV}}{\text{Ra Th}} \xrightarrow{\alpha} \underset{\text{II}}{\text{Th X}} \xrightarrow{\alpha} \underset{\text{O}}{\text{Em}} \xrightarrow{\alpha} \underset{\text{VI}}{\text{A}} \xrightarrow{\alpha} \underset{\text{IV}}{\text{B}} \xrightarrow{\beta} \underset{\text{V}}{\text{C}} \rightarrow$$

The numerals underneath represent the number of the column in the periodic table (or the value of the electropositive valency). The character of the change is indicated above the arrows.

But the matter goes further than this. In some cases an α ray change is followed by two successive changes in which an α ray is not expelled. The last element then occupies the same column in the periodic table as the original element. In such cases these two elements have, so far as can be ascertained, identical chemical properties, and are incapable of being separated by chemical methods. There is evidence[‡] also that their emission spectra are identical. These results are surprising at first sight since the atomic weights of the elements in question must differ by the atomic weight of helium approximately, or about four units. In contrast to their chemical properties, the radioactive properties of two such elements are quite different. These phenomena receive a plausible explanation on the view, advocated by Rutherford, that the atoms are built up of electrons revolving round a massive central nucleus of small dimensions, if we adopt the hypothesis that both the α and β particles are ejected from the nucleus itself. For the chemical and spectroscopic properties of such an atom will be determined almost entirely by the charge of the nucleus and hardly at all by its mass. The ejection of an α particle, which is a positively charged helium atom with twice the electronic charge, will diminish the total positive valency of the nucleus by two, and the ejection of an electron will increase it by one. The resulting nuclei will be those appropriate to atoms whose positions in the periodic table are those actually found.

---

[*] F. Soddy, *Chem. News*, vol. cvii. p. 97 (1913); G. v. Hevesy, *Phys. Zeits.* vol. xiv. p. 49 (1913); K. Fajans, *Phys. Zeits.* vol. xiv. pp. 131, 136 (1913); *Verh. d. Deutsch. Phys. Ges.* vol. xv. p. 240 (1913).

[†] *Chem. News*, vol. cvii. p. 49 (1913).

[‡] Russell and Rossi, *Roy. Soc. Proc.* A, vol. lxxxvii. p. 478 (1912).

This position is strengthened by a number of other lines of evidence. Thus J. J. Thomson* has discovered that neon consists of two gaseous elements having the same spectrum and chemical properties but differing in atomic weight. We have seen also that the X-ray spectra of the elements are not accurately functions of the atomic weight but are accurately determined by successive whole numbers. The sequence of these numbers is that of the atoms in order of ascending atomic weight except for the elements whose positions in the periodic table are anomalous. It seems natural to interpret the "atomic numbers" as measures of the charge carried by the positively charged nucleus. These facts do not receive any obvious explanation on the Kelvin-Thomson theory of the atom.

Bohr's Theory of the Behaviour of Electrons in Atoms.

This theory†, which frankly discards dynamical principles, has already been referred to in Chap. XVI. As its chief success so far has been in the explanation of the frequency of spectral lines and other phenomena which depend intimately on atomic structure, it may be permissible to recapitulate here the chief assumptions on which it rests. It is based on the idea of negative electrons revolving round a massive positive nucleus of negligibly small dimensions, the mutual force following the electrostatic law of the inverse square. Stability and definiteness of frequency, which, as we have seen, are dynamically impossible with such a structure, are secured by the introduction of the two following hypotheses:

---

* *Roy. Soc. Proc.* A, vol. LXXXIX. p. 1 (1913).

† N. Bohr, *Phil. Mag.* vol. XXVI. pp. 1, 476, 857 (1913), vol. XXVII. p. 506 (1914).

(1) That the stable motions are those for which the angular momentum of the electrons is an integral multiple of $h/2\pi$, and

(2) That radiation is emitted when there is a change from one possible configuration of the electrons to another. The radiation thus emitted is assumed to be monochromatic and its frequency ν to be determined by the equation $W_1 - W_2 = h\nu$, where h is Planck's constant and W_1 and W_2 are the energies of the system in the two configurations under consideration.

These assumptions lead immediately to the following formula for the possible frequencies which may be emitted by the simplest atom (that which contains a single electron when neutral):

$$\nu = \frac{2\pi^2 me^4}{h^3}\left(\frac{1}{\tau_2^2} - \frac{1}{\tau_1^2}\right) \quad \text{..................(29),}$$

where τ_2 and τ_1 are whole numbers. When $\tau_2 = 2$ and τ_1 takes successive integral values this gives Balmer's series. The strong point in favour of this theory is that it determines the absolute constant in the spectral series in terms of the fundamental constants m, e and h, and the agreement is exact to the accuracy within which these constants are known. The spectroscopic value of the constant K in Balmer's series is $3{\cdot}290 \times 10^{15}$. When the best values of e, m and h are used the calculated value of the factor $\frac{2\pi^2 me^4}{h^3}$ in (29) is $3{\cdot}26 \times 10^{15}$. By putting $\tau_2 = 3$ another known hydrogen series is obtained.

When the central charge has twice the value just considered and there is a single electron revolving round it, we should expect to get the spectrum of a helium atom which had lost both of its electrons (an α particle) and subsequently combined with a single electron. As a matter of fact the series calculated for this system include Pickering's* series and certain series recently observed by Fowler†. None of these series have been obtained in the laboratory except in the presence of helium.

---

* *Astrophys. Journ.* vol. IV. p. 369 (1896), vol. V. p. 92 (1897).

† *Month. Not. Roy. Astr. Soc.* vol. LXXIII. (1912).

CHAPTER XXII

GRAVITATION

General Characteristics.

THE position of physical theories of gravitation to-day is almost as speculative as it has been ever since the days of Newton. Nevertheless we cannot afford entirely to overlook the bearing of the electron theory on gravitational action, particularly as we have found the theory to be capable of giving a fair account of the other known physical phenomena. Paradoxical though it may seem, it is quite likely that one of the chief difficulties in the way of a physical theory of gravitation lies in the extreme simplicity of the known laws of gravitational action. The prospect of addition to our knowledge of gravitation by experiment is not very hopeful on account of the smallness of the forces concerned. These are large enough, of course, when we are dealing with the enormous aggregations of matter familiar to astronomy, but they are exceedingly small with masses which can be controlled in the laboratory.

The fundamental law of gravitation, which was enunciated by Newton, may be expressed by means of the equation

$$F = k\frac{dm\,dm'}{r^2},$$

where F is the attractive force between two material particles whose masses are dm and dm' at a distance r apart, and k is the gravitational constant which is equal to $6{\cdot}6 \times 10^{-8}$ in C.G.S. units. This inverse square law of attraction is the foundation of the astronomical calculations of the orbits of the heavenly bodies. As these are very exact it follows that the index of r is 2 with great precision.

occur. In this case the functions F will all be constants characteristic of the electrons. It also follows by symmetry that $F(E_2e_1) \equiv F(E_1e_2)$. Since all the functions are very nearly equal to unity, we may write

$$F(E_1E_2) = 1 + a, \quad F(e_1e_2) = 1 + b, \quad F(E_1e_2) = F(E_2e_1) = 1 + c,$$

where a, b and c are very small constant quantities. The repulsion between two neutral particles is therefore equal, on this supposition, to $\dfrac{N_1N_2E^2}{4\pi r^2}(a + b - 2c)$. This will be negative if $2c$ is greater than $a + b$, so that in order to account for gravitational attraction it is necessary to suppose that the attraction between two unlike charges is slightly greater than the repulsion between two like charges.

On the further hypothesis that all the mass of the substances is electromagnetic it will be proportional to the numbers N_1 and N_2 of the electrons which the particles contain, except in so far as the electromagnetic mass of an electron may vary somewhat in different atoms. If M_1 and M_2 are the masses of the particles at P and Q respectively and if M is the mass of a positive and m that of a negative electron, then $N_1 = \dfrac{M_1}{M + m}$ and $N_2 = \dfrac{M_2}{M + m}$, so that the repulsion becomes

$$\frac{M_1M_2}{4\pi r^2}\left\{\frac{E^2}{(M + m)^2}(a + b - 2c)\right\} \quad \text{..............(3).}$$

Since the quantity in brackets is a universal constant, subject to the hypotheses which we have made, the attraction accords with the Newtonian law.

The preceding result is based on the hypothesis that the mass of the electrons is not appreciably different in different substances. We shall see in the next paragraph that we cannot be sure that this is the case.

The Atomic Weights of the Elements.

It is now fairly certain that the number of negative electrons in any atom is equal to about half the atomic weight or, at all events, is comparable with the atomic weight. As the value of e/m for the negative electrons is very large it follows that only a very small proportion of the mass of an atom can arise from the electromagnetic mass of the negative electrons. If the mass of the atoms is of electromagnetic origin it must be associated almost entirely with the positive electricity. As we have seen, there is a considerable body of evidence in favour of the view that the positive electricity is confined to a very minute region of the atom, so that practically the whole of the atomic mass will reside in this nucleus. The estimated linear dimensions of the nucleus are very small (about the same as the linear dimensions of the negative electron) so that considerable overlapping of the fields of the constituent positive electrons may be expected. We should therefore expect the electromagnetic mass of such a nucleus not to be equal to the sum of the masses of the constituent positive electrons on account of the mutual interference of their electric fields. It is impossible to say precisely what difference would be expected to arise in this way, but it is likely to be only a small fraction of the total mass.

The expulsion of α rays from the atoms of the radioactive elements shows that the nucleus is made up of separable parts. However, the α particles are atoms of helium and have very nearly four times the mass of the lightest known positively charged particle, the positive hydrogen ion. Thus unless we accept the hypothesis that there is more than one kind of positive electron we cannot regard the α particle as such an ultimate unit. If we take the hydrogen ion to be the positive electron, the value of e/m which it possesses shows that its linear dimensions are about 10^{-4} of that of the nuclei. If such positive electrons were distributed fairly uniformly throughout the volume of the nucleus the overlapping of the fields would have little influence on the electromagnetic mass. The electromagnetic mass of an atom would then be very close to the sum of the masses of the positive electrons, and the truth of the statement at the end of the preceding paragraph would follow. We cannot, however, be sure of this deduction as we are completely ignorant of the constitution of the hypothetical nucleus.

It might be thought that further information might be obtainable from a consideration of the atomic weights of the elements. If the mass of an atom can be treated with sufficient accuracy as the sum of the masses of the constituent positive and negative electrons and if each kind of atom can be regarded simply as a neutral combination of a given number of such units then, to the same degree of accuracy, the atomic weights of the different elements will be proportional to whole numbers. This is on the supposition that there is only one kind of positive and one kind of negative electron. If there are n different kinds of electrons of masses m_1, m_2, etc., then the masses of the atoms will be of the form $a_1 m_1 + a_2 m_2 + \ldots$, where a_1, a_2, etc. are integral numbers starting from zero. It is well known that the atomic weights of the elements exhibit an approximation to whole numbers which cannot be fortuitous*; but the deviations from integrals, or from the linear law just referred to, cannot now be regarded as likely to arise from modifications of the electromagnetic mass of the individual electrons. The work of Soddy and others (see p. 583) on the chemical properties of the radioactive elements has shown that elementary bodies having identical chemical properties and different atomic weights frequently coexist. It is therefore likely that a great many so-called chemical elements are mixtures of atoms having the same chemical properties but differing in mass. The experimentally determined atomic weight is, on such a view, the average of the weights of the constituent atoms and depends on the relative proportion in which they are present.

There is another point which deserves consideration in this connection. The two lightest positively charged particles with which we are familiar are the hydrogen ion and the α particle. According to the views of Rutherford and Bohr both of these particles consist only of positive electricity. The α particle has twice the actual charge and approximately half the specific charge (e/m) of the hydrogen ion. It is quite possible that the α particle is made up of four hydrogen ions cemented into a nucleus by means of two negative electrons, but it has not, so far, been shown that the α particle can be broken up into anything smaller; so that we shall examine tentatively the consequences of supposing that it is a fundamental element of positive electricity independent of the hydrogen ion. In that case we have to deal with two positive electrons, one of which has associated with a given charge approximately twice as much mass as the other. On p. 591 we considered gravitational attraction from the point of view that it arose from a minute lack of compensation between the attractions and repulsions of the ultimate electric charges. By supposing the way of describing very much the same position is to say that the electrostatic law of force represents what would occur if the infinitesimal elements of charge were at rest in the undisturbed aether, that is to say in aether devoid of energy. The classic electrostatic law of force would then represent exactly what would occur if the aether were homogeneous. But in the immediate neighbourhood of an actual element of electric charge there is an immense disturbance going on and, as we have seen, the mass of the element is a measure of the energy of this disturbance. We might therefore expect that the field of force emanating from such a region would be modified to an extent depending on the local disturbance; and also that the reaction, to an external field, of an element of charge thus situated would be similarly affected. Under these circumstances the force between two positive electrons at a given distance apart will not necessarily be equal to the force between two negative electrons of equal charge separated by the same distance, on account of the difference of the masses; neither will the attraction between unlike electrons necessarily be equal to the repulsion between like electrons carrying equal charges.

---

* Strutt, *Phil. Mag.* [6] vol. I. p. 311 (1901).

Unfortunately there is an apparently fatal objection to any view which makes the modified law of force depend on what is practically nothing more than an alteration in the effective charge produced by the local energy or mass. For the law of force will then involve the product of two factors each of which depends on the state of only one of the two separate elements. If dm denotes the mass which is associated with the element of charge de, we can take $\frac{\partial m}{\partial e}$ as a measure of the intensity of the local disturbance. Using Taylor's Theorem we see that the force between two elements ∂E and $\partial e'$ instead of being equal to $\frac{\partial E \partial e'}{4\pi r^2}$ will be

$$\frac{\partial E \partial e'}{4\pi r^2}\left\{1 + a\frac{\partial m'}{\partial e'} + A\frac{\partial M}{\partial E} + aA\frac{\partial m'}{\partial e'}\frac{\partial M}{\partial E} + \dots\right\} \dots\dots\dots (4),$$

where A and a are fundamental constants which must be equal, by symmetry, when the charges are of like sign; otherwise this is not necessarily the case. Keeping the capital letters for positively charged elements and the small letters for negative charges it

If the condition for neutrality is $\partial E + \partial e = \partial E' + \partial e' = 0$, the force between neutral systems is

$$g\,(\partial M + \partial m)\,(\partial M' + \partial m')/4\pi r^2 \quad \ldots\ldots\ldots\ldots\ldots(9).$$

As g may be negative, this agrees with the Newtonian law.

According to formula (8) the terms involving a and c might make an important difference to the force exerted on an electron by a gravitational field, but would have no effect on the attraction of two neutral particles, which is determined entirely by the term containing g. The introduction of this term is practically equivalent to attributing the Newtonian law of mass attraction to each pair of elements of charge however small.

Conditions for Neutrality.

The introduction of gravitational forces complicates electrostatic problems very considerably. The condition for electrical equilibrium in any conductor is that the stream of positive electricity in any direction at any point should be equal to the stream of negative electricity in the same direction at that point. In general this condition may be very complicated. If we may suppose that only the negative electrons are free to move, as the positive current is zero the negative current must be zero also. If we are also free to neglect effects arising from differences of pressure of the electrons at different points, the condition will be satisfied if at every point the electric force on an electron is exactly balanced by the gravitational force, the resultant average force on an electron then being zero. There will thus be a potential gradient in a conductor in equilibrium in a gravitational field; this potential gradient is, however, too small to be detected experimentally.

The way to define a neutral material particle is not an obvious matter under these circumstances. We might, for example, define a neutral material particle as one which exerted no force on a negative electron; but equally valid reasons could be urged for making the vanishing of the force on a positive electron the condition for neutrality. If both these conditions were satisfied there would be no force between two neutral particles, so that they are not compatible with the existence of gravitational attraction. In accordance with the considerations of the last section we might define a neutral particle as one which would give rise to forces exerting system are referred to the set of axes which we have associated with that system. They are clearly

$$\frac{\partial^2 x}{\partial t^2}=0, \quad \frac{\partial^2 y}{\partial t^2}=0, \quad \frac{\partial^2 z}{\partial t^2}=-\gamma.$$

Now let us assume that our two systems are physically identical, i.e. that it is impossible to distinguish between them by any physical means whatever. This is tantamount to denying absolute acceleration in the same way that the principle of relativity denies absolute velocity. The equivalence of the two systems is obvious if we confine ourselves to the Newtonian mechanics. This is not the case however when we turn to some of the results which we have deduced from the principle of relativity.

It is shown on p. 318 that the mass of any system involves a part E/c^2, where E is the radiant electromagnetic energy of the system and c is the velocity of light. E/c^2 is mass in the sense of a coefficient of inertia but we have not proved that it is subject to gravitational attraction, in the way that the ordinary mass of the system is. If this mass is not subject to gravitational attraction it is evident that the equivalence of the systems κ and κ' will not hold exactly; systems in a uniform gravitational field will only fall uniformly provided the inertia of their electromagnetic energy may be neglected. A formal proof that if κ and κ' are exactly equivalent it is necessary that E/c^2 should represent gravitating mass, as well as inertia, is easily constructed

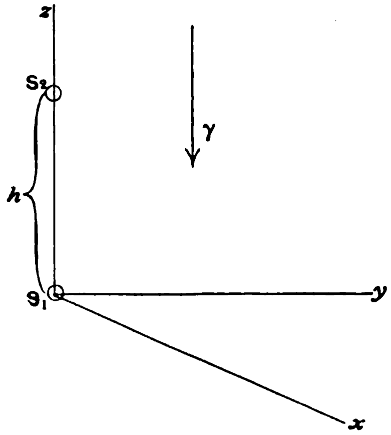


Fig. 58.

Let the diagram represent the axes κ in the field of uniform gravitational force. Let S_2, S_1 represent two sufficiently minute material systems at a distance h apart along the z axis. If S_2 and S_1 are small enough, h will remain invariable. The systems will simply slide along the z axis with uniform acceleration, as though they were held apart by a rigid connection of length h. Suppose that S_2 is allowed to send a certain amount of radiant energy E which is received by S_1. The energy is measured at S_2 and S_1 by instruments which can be brought together and compared. As we do not yet know anything about the effect of gravitation on electromagnetic energy we are unable to say anything about what takes place during the transference. We can, however, find out a great deal about it if we admit that the system κ' is equivalent to the system κ. For in the system κ' let us measure the energy of the radiation with reference to a set of axes κ_0 that are at rest with reference to κ' at the instant at which the radiation leaves S_2. If the energy thus measured is E when it leaves S_2, it will, by the theory of relativity*, be equal to $E(1 + \gamma h/c^2)$ by the time it is at S_1. For S_1 is then moving with reference to the axes κ_0 with the relative velocity $v = \gamma t = \gamma h/c$. This result is true only to the first order of the small quantity $\gamma h/c^2$.

On the hypothesis that the two systems are equivalent, exactly the same result will have to hold with reference to the system of axes κ in the gravitating system. In this case γh is equal to ϕ, the difference in the gravitational potential between the points S_2 and S_1; so that if E_2 and E_1 are the energies of a given quantity of radiation at S_2 and S_1 respectively, then

$$E_1 = E_2(1 + \phi/c^2).$$

If we are to retain the principle of the conservation of energy in these cases it follows that work equal to $E_2\phi/c^2$ is done on the radiant energy by the gravitational forces. This is the same as if the radiant energy possessed gravitational mass equal to E/c^2. We therefore conclude that E/c^2, which the principle of relativity shows to be the inert mass of radiation, represents its gravitating mass also.

---

* Einstein, *Ann. der Physik* vol. XVII. p. 913 (1905).

The Velocity of Light in the Gravitational Field.

Returning to the system κ', let ν_2 be the frequency of the light emitted by S_2; then by Doppler's principle the light received by S_1 will appear to have a different frequency, viz. $\nu_1=\nu_2(1+\gamma h/c^2)$ to the first approximation. If the systems κ and κ' are equivalent, the gravitational field will therefore produce a change in the frequency ν_2 of the radiation which leaves S_2 to the value

$$\nu_1=\nu_2(1+\phi/c^2)$$

at a place S_1 where the gravitational potential differs by ϕ from its value at S_2. These results are true when the clocks used to measure time at S_2 and S_1 agree with one another when brought together. But the distance h is constant in the κ system; so that, if the radiation were being continually emitted by S_1 and received by S_2, there would be a continuous accumulation or depletion of waves in the space between them, if ν_1 and ν_2 were unequal. This contradiction indicates that the times at different parts of a gravitational field are not correctly given by clocks which agree when brought together. To give times which do not lead to a contradiction the clock at S_2 must go $1+\phi/c^2$ times slower than the clock at S_1 when they are brought together and compared. Measured with such clocks the frequencies at S_2 and S_1 become the same and the number of waves in the stretch h is independent of the lapse of time, when the emission and reception are steady and continuous.

Now if we measure the velocity of light in different places of the system κ' with clocks which agree when compared together we get always the constant quantity c. If κ and κ' are equivalent, the same is true of the gravitating system κ. But these clocks do not go at the right rate in different parts of κ; whence it follows that the velocity of light is not the same at different parts of κ but varies with the gravitational potential according to the relation

$$c_1=c_2\left(1+\frac{\phi}{c_1^2}\right)$$

Thus the principle of the constancy of the velocity of light which Einstein made the basis of the principle of relativity does not hold for gravitational fields according to this theory.

It is easy to calculate the curvature of a ray of light produced by gravitation on this view. Take two points on the wave front at distance dl apart. At time dt later the wave front will be tangential to spheres of radii cdt and $\left(c + \frac{\partial c}{\partial l}\,dl\right)dt$ described about the two points. The rotation $\delta\theta$ of the wave front in time δt is thus

$$\delta\theta = \frac{\partial c}{\partial l}\,\delta t.$$

If δn is an element of the normal to the wave front $cdt = \delta n$,

$$\frac{\partial\theta}{\partial n} = \frac{1}{c}\frac{\partial c}{\partial l} = \frac{1}{c^2}\frac{\partial\phi}{\partial l}$$

to a sufficient approximation.

Let us apply this to the case of light passing within a distance δ of the centre of a star of mass m. Let k be the gravitational constant, r the instantaneous distance of the ray of light from the centre of the star and ψ the angle between r and the perpendicular from the star on the straight path of the undeflected light. Then $\phi = km/r$ and the small total angular deflexion $\delta\theta$ is

$$\delta\theta = \frac{1}{c^2}\int\frac{\partial\phi}{\partial l}\,dn = -\frac{km}{c^2}\int_{-\pi/2}^{\pi/2}\frac{1}{r}\frac{\partial r}{\partial l}\frac{d\psi}{\cos\psi}$$

$$= -\frac{km}{c^2\delta}\int_{-\pi/2}^{\pi/2}\cos\psi\,d\psi = -\frac{2km}{c^2\delta}.$$

The deflexion is directed towards the star, so that if the light comes from a second star the apparent angular distance between their centres will be increased; in the case of the light from a star passing just outside the sun's disc it amounts to 0·83 second of arc.

In a later paper Einstein* has extended the investigation so as to cover the relation between gravitation and electromagnetic phenomena other than mass and the velocity of radiation, using similar methods. The smallness of the expected effects appears to preclude the possibility of their experimental verification in the near future.

---

* *Ann. der Physik*, vol. XXXVIII. p. 355 (1912).

INDEX OF SUBJECTS

α rays, scattering of 490
Aberration of light 269, 306
Absorbing media 163
Absorption, feeble 166; of characteristic X rays 497 ff.
Absorption of light 144; coefficient of 165; near critical frequencies 178
Acceleration and force 260
Aether 268; and the principle of relativity 323
Aether pulse theory of the Roentgen rays 483
Airy's experiment 271
Alloys, conductivity of 465
Arago's experiment 271
Atom, definiteness of 549; of Kelvin and Thomson 554
Atomic, numbers 510, 583; structure 491, 494, 549, 554, 564; theory, application to the 6; weights of elements, and gravitational attraction 592; weights of radioactive elements 583
Atoms, Bohr's theory of 395, 585; number of electrons in 488, 494, 495, 564
Attraction, gravitational, compared with electrostatic 590
Avogadro's number 6
Axes, moving 286, 297

β rays, deflexion by electric field 236; deflexion by magnetic field 237, 240; scattering of 490
β secondary X rays 503
Band spectra 517
Black body radiation 330, 433
Bohr's theory of atoms 395, 585; theory of spectra 585
Bucherer's experiment 239

Characteristic X rays, absorption of 497 ff.
Charge, accelerated, field due to 247, 249; accelerated, radiation from 257; central, of atom 494, 510, 583, 585; electric, with variable velocity 243; on electron, determination of 3, 356
Charged system, in uniform motion 217
Charges, fictitious 57
Chemical combination 570; and transference of electrons 571, 574, 580
Collisions between particles under central forces 417
Combination, chemical 570
Conditions at boundary between two dielectric media 45
Conducting electrons 463
Conducting media, propagation of electromagnetic waves in 135; reflexion at boundary of 137
Conduction in metals, electron theory of 408, 413, 463; simple theory of 409
Conductivity, electric, of alloys 465; electric, of bad conductors 467; electrical 410, 420, 422; electrical, for periodic forces 430; of metals for heat 411, 421
Constant, Boltzmann's 406; Planck's 356, 469, 506, 509, 585
Contact difference of potential 454
Contractile electron 232
Contraction, Fitzgerald's 280, 291
Correlation, principle of 286, 290
Corresponding states, of magnetization 382
Coulomb's Law 23
Critical temperature, magnetic 389
Crystals and X rays 507
Curie's Law 378
Current, convection 98; displacement 97; electric 96; force on an element of, in magnetic field 114; magnetic force due to element of electric 113; polarization 98
Currents, induction of 100

Diamagnetism 361, 368
Dielectric constant, electron theory of 74
Dielectric media 39; Poisson's theory of 47; electron theory of 58
Differential equations satisfied by vectors when charges are present 186

Mass, electromagnetic, of rigid electron 228; electromagnetic, variation with velocity 234; longitudinal and transverse 229 ff.; variation with velocity on the principle of relativity 312
Material particles, conditions for neutrality 598; forces between 595, 599
Maxwell's stresses 35, 81
Mechanics of electron 310
Michelson and Morley's experiment 277
Moment, magnetic, of revolving electron 367
Momentum, and radiation pressure 212; effect of inertia of energy on 320; electromagnetic 208, of charge in uniform motion 221, of isolated system 215, of rigid electron 227; of moving particle 314, 319
Moving axes 286, 297

Natural rotation of plane of polarization 545
Negative electron 7
Neutral material particle 598
Newton's law of gravitation 588
Number of electrons in atoms 488, 494, 495, 564
Numbers, atomic 510, 583

Optical, convection, Stokes's theory of 275; effects due to motion of refracting medium 273
Oscillator, energy of 349

Paramagnetic gas 379
Paramagnetism 361, 376, 379
Peltier effect 429, 455, 457
Periodic law, and radioactive elements 583; and structure of atoms 565
Permeability, diamagnetic 371
Phosphorescence 530
Photoelectric action 469
Planck's radiation formula 347, 354, 355
Point charge in uniform motion, field due to 248
Poisson's equation 26, 42; solution of equations of propagation 118; theory of dielectric media 47
Polarization, electric 51; electron theory of 64, 70; magnetic rotation of plane of 535, 542; natural rotation of plane of 545; of X rays 488; potential due to 60; variable 54
Positive electricity, structure of 580
Positive electron 8, 482, 580, 593
Potential, at internal points 19; contact difference of 454; electric 15; magnetic, due to electric current 86; propagated 193; scalar 194; vector 194, 197
Poynting's theorem 203
Pressure of radiation 209
Pressure shift of spectral lines 533
Principle of correlation 286, 290
Principle of least time 271
Principle of relativity 296, 322
Probability and entropy 400
Probability, of distribution of energy among oscillators 352, 357; of statistical distribution of a gas 401
Propagation, equations of 115
Pulse theory of Roentgen rays 483
Pyrrhotite, magnetic properties of 384

Quasi-stationary motion 262

Radiant energy, gravitational mass of 603
Radiation, and temperature 326, 433; complete 328, 433; energy density of complete, in refracting medium 332; formula of Planck 347, 355; formula of Rayleigh and Jeans 343; formula of Wien 354; from accelerated electric charge 257; isotropic 211; pressure and momentum 212; pressure of 209; reaction of, on accelerated charge 263; reflexion of, at a moving mirror 277, 335; secondary, asymmetrical emission of 477 ff.; types of 476
Radioactive elements and periodic law 583
Ratio of electromagnetic to electrostatic unit 122
Rayleigh and Jeans's radiation formula 343
Rays, residual 157
Reaction of radiation on moving charge 263
Reflecting power, of metals 140; of quartz 159
Reflexion, change of phase on 138, 168; of electromagnetic waves 129; of electrons by conductors 468; of light by moving mirror 277, 335; of light near critical frequencies 178
Refracting medium, optical effects due to motion of 273; theory of propagation of light in moving 284
Refraction, of compounds and mixtures 150; of electromagnetic waves 129
Refractive index 149, 165
Relative theory of gravitation 601
Relativity, principle of 296; dependence of physical quantities on velocity 322; experimental test of 242, 315, 318
Residual rays from absorbing media 167
Resistance, change in a magnetic field 439
Reststrahlen 159
Roentgen rays 482 ff., 496 ff., 507 ff.

Scattering, of α and β rays 490; of α and β rays, Rutherford's theory 494; of α and β rays, Thomson's theory 491; of X rays 485

Secondary rays 477, 496
Series in band spectra 517
Series, spectral 513, 585
Shell, force on a magnetic 83
Shells, polarized 62
Specific heat of electricity 429, 448, 451, 452
Specific heats, at low temperatures 358; Debye's theory 360; Einstein's theory 357; of ferromagnetic substances 397
Spectra, band 517
Spectra, Bohr's theory of 585; X ray of elements 510
Spectral emission, theories of 525, 585
Spectral lines, definiteness of frequency of 551; displacement in electric field 534; effect of pressure on 533
Spectral series 513, 585
Spectroscopic phenomena 512, 517, 525, 551, 585
Stability, of electrons in atom 552, 554, 558, 562
Stark effect 534
Stefan's law 333
Stirling's approximation 402
Stokes's theorem 91; theory of optical convection 275
Stresses, Maxwell's 35, 81, 207 ff.
Structural formulae, in chemistry 575
Structure of positive electricity 580

Temperature and radiation 326, 433
Thermionics 441
Thermodynamics and kinetic theory 399
Thermoelectricity 425, 448, 455, 459, 461
Thermoelectromotive force 459
Thermomagnetic effects 434
Thomson effect 429, 448, 452
Transparency of quartz 159
Tubes of force 16, 22

Unit of current 89
Unit of electric charge 13
Units, electrical 109

Valency, chemical, and atomic structure 566, 573
Values, actual and mean 65
Velocity, addition of, on principle of relativity 302; of light in the gravitational field 604; of propagation of electro-magnetic field 121; of secondary β rays emitted by characteristic X rays 504; wave 252

Wave of reorganization 258
Wave, plane-polarized electromagnetic 124
Wave-length of X rays 510
Waves, electromagnetic 115
Wien's law 337; displacement law 343; radiation formula 354

X ray spectra 510
X rays, and crystals 507; β secondary 503; characteristic, absorption of 497 ff.; characteristic, velocity of secondary β rays emitted by 504; polarization of 488; scattering of 485; wave-length of 510

Zeeman effect 373, 518, 523; inverse 535; theory of 373, 518

CAMBRIDGE: PRINTED BY JOHN CLAY, M.A. AT THE UNIVERSITY PRESS